

# Review and Perspectives on Human Emotion for Connected Automated Vehicles

Wenbo Li<sup>1</sup> · Guofa Li<sup>2</sup> · Ruichen Tan<sup>5</sup> · Cong Wang<sup>1</sup> · Zemin Sun<sup>3</sup> · Ying Li<sup>4</sup> · Gang Guo<sup>2</sup> · Dongpu Cao<sup>1</sup> · Keqiang Li<sup>1</sup>

Received: 28 September 2022 / Accepted: 1 September 2023 / Published online: 30 January 2024  
© The Author(s) 2024

## Abstract

The progression toward automated driving and the latest advancement in vehicular networking have led to novel and natural human-vehicle-road systems, in which affective human-vehicle interaction is a crucial factor affecting the acceptance, safety, comfort, and traffic efficiency of connected and automated vehicles (CAVs). This development has inspired increasing interest in how to develop affective interaction framework for intelligent cockpit in CAVs. To enable affective human-vehicle interactions in CAVs, knowledge from multiple research areas is needed, including automotive engineering, transportation engineering, human-machine interaction, computer science, communication, as well as industrial engineering. However, there is currently no systematic survey considering the close relationship between human-vehicle-road and human emotion in the human-vehicle-road coupling process in the CAV context. To facilitate progress in this area, this paper provides a comprehensive literature survey on emotion-related studies from multi-aspects for better design of affective interaction in intelligent cockpit for CAVs. This paper discusses the multimodal expression of human emotions, investigates the human emotion experiment in driving, and particularly emphasizes previous knowledge on human emotion detection, regulation, as well as their applications in CAVs. The promising research perspectives are outlined for researchers and engineers from different research areas to develop CAVs with better acceptance, safety, comfort, and enjoyment for users.

**Keywords** Intelligent vehicles · Intelligent cockpit · Human-machine interaction · Emotion recognition · Emotion regulation

## Abbreviations

ACAD	Affective connected automated driving	CCC	Consistency correlation coefficient
ALE	Activation likelihood estimation	CNS	Central nervous system
ANS	Autonomic nervous system	DES	Differential emotions scale
AUs	Action units	ECG	Electrocardiography
BVP	Blood volume pulse	EDA	Electrodermal activity
CAVS	Connected and automated vehicles	EEG	Electroencephalography
		EMG	Electromyography
		FACS	Facial motion coding system
		fNIRs	Functional near-infrared spectroscopy
		HR	Heart rate
		HRV	Heart rate variability
		JAFFE	Japanese female facial expression
		MSE	Mean squared error
		NIR	Near-infrared
		PANAS	Positive and negative affect schedule
		PPG	Photoplethysmography
		RESP	Respiration
		RF	Radio frequency
		SAM	Self-assessment manikin
		SCL	Skin conductance level
		ST	Skin temperature

✉ Wenbo Li  
wenboli@tsinghua.edu.cn

<sup>1</sup> School of Vehicle and Mobility, Tsinghua University, 30 Shuangqing Rd, Beijing 100084, China

<sup>2</sup> College of Mechanical and Vehicle Engineering, Chongqing University, 174 Shazheng Rd, Chongqing 400044, China

<sup>3</sup> College of Computer Science and Technology, Jilin University, 2699 Qianjin Rd, Changchun 130012, China

<sup>4</sup> School of Mechanical Engineering, Beijing Institute of Technology, 5 Zhongguancun South Rd, Beijing 100081, China

<sup>5</sup> Lyles School of Civil Engineering, Purdue University, 550 Stadium Mall Drive, West Lafayette, IN 47907, USA

V2C	Vehicle-to-cloud
V2I	Vehicle-to-infrastructure
V2P	Vehicle-to-pedestrian
V2V	Vehicle-to-vehicle

## 1 Introduction

With the advancements in artificial intelligence and informatics communication technology, connected and automated vehicles (CAVs) are experiencing rapid development globally in terms of hardware, software, systems engineering, infrastructure, and policy [1, 2]. CAVs can use sensors, controllers, actuators, and integrate communication and network technologies to perform environmental perception, intelligent decision-making, and collaborative control, enabling safe, efficient, comfortable, and eco-friendly automated driving [3, 4]. Thus, the CAV has become an intelligent mobile terminal carrying rich functions and services. In the foreseeable future, as a part of the CAV that users can directly experience, intelligent cockpit will provide drivers and occupants with a safer, more efficient and more enjoyable ride experience.

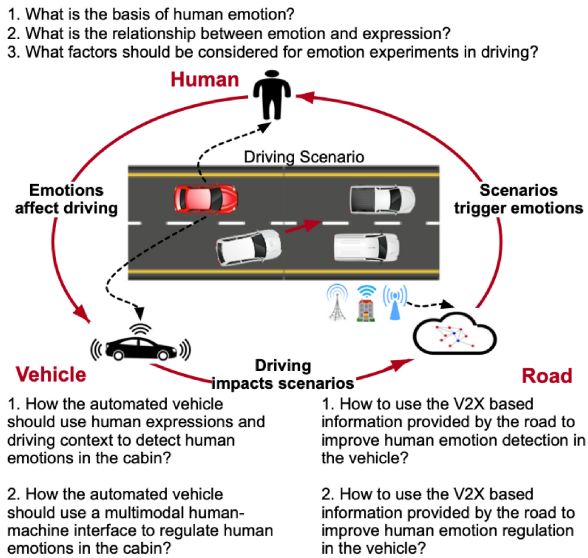
Intelligent cockpit in CAVs is a mobile space with humans at its core, capable of dynamically perceiving and understanding human behavior and providing feedback accordingly [5]. The development of the intelligent cockpit will expand and deepen the relationship between humans, vehicles, and roads, opening up new opportunities for research and innovation. Among these opportunities, affective interaction in the intelligent cockpit is particularly valuable for improving the acceptance, safety, comfort, and enjoyment of human behavior and emotions in the context of CAVs [6–8]. Affective interaction refers to using sensors and other equipment to enable the intelligent cockpit to perceive human emotional expressions, accurately identify human emotions, and make corresponding decisions based on the recognition results. The intelligent cockpit then provides related functions or services to humans through different interaction modes, according to the decision-making strategy.

The role of affective interaction within intelligent cockpits plays an essential role in the development of CAVs from driver assistance to automated driving, from single-vehicle intelligent driving to multi-vehicle collaborative driving [9]. For single-vehicle automated driving, due to the trust concerns, human acceptance of the CAV can still be a significant issue for the popularities of CAVs, while the human emotional state can be a critical index for both real-time and

long-term human trust [10, 11]. Besides, driving context can frequently change between a potentially stressful environment (requiring a high degree of concentration and responsibility to the human) and a relaxed environment. Affective interaction with the occupants will augment their driving experience and comfort, especially when sharing the driving authority with the CAV [12]. For multi-vehicle collaborative driving, in the development process of CAVs, there is bound to be a long-term mixed traffic situation where automated driving and human driving coexist. Studying the expression and behavior of human emotions and detecting and regulating emotions will help CAVs improve road safety and traffic efficiency [9].

While affective interaction could not only provide great opportunities in improving human acceptance, safety, and comfort of CAVs, and but also bring enormous potential to enhance the safety and mobility of public transport, there are still challenges for future research. First, developing a proper affective interaction framework for the intelligent cockpit in CAVs requires cross-disciplinary knowledge, including automotive engineering, cognitive psychology, human factor, affective computing, and intelligent transportation system. Collaboration between experts in these fields is necessary to develop a comprehensive and efficacious affective interaction system [13, 14]. More importantly, in CAVs, the relationship between the human, vehicle and road is more complicated than in traditional vehicles, as the road scenario causes the generation of human emotions in driving, and the human emotional responses will continue to impact the road scenario. Therefore, studying human-vehicle-road coupling is crucial for developing an efficient affective interaction system in CAVs. As illustrated in Fig. 1 shows the coupling delineates the multifaceted relationship between the human operator, the vehicle, and the surrounding road environment in CAVs. Different road scenarios, such as urban and rural, can trigger human emotions and behavior in driving. Human emotions such as stress and anxiety can affect human behavior and driving performance, as well as the perception of the road environment and the vehicle's behavior. The vehicle's acceleration, deceleration, speed, and other parameters can affect road scenarios.

Several studies have reviewed driver emotion in terms of driver emotion recognition [15] and regulation techniques [16]. However, very few consider the intricate interplay between humans, vehicles, roads, and the resulting emotional states within the context of CAVs. Besides, fundamental understandings and potential applications of human emotions in the context of CAVs are rarely considered. Considering these limitations, in this study,



**Fig. 1** The system description of the human-vehicle-road in CAVs, and the topics involved in each part (V2X = vehicular communication to everything)

a literature review is presented to analyze human-vehicle-road affective interactions for future CAVs. In sum, the works of this study can be summarized as follows. To provide clarity, the organizational structure of this study, along with pertinent research domains, is depicted in Fig. 2.

- First, the basic concepts of human emotions are introduced to have a clear understanding of the production process and classification of human emotions.
- Second, the relationship between different emotions and emotional expressions is summarized. Facial expressions, speech expressions, body gesture expressions, and physi-

ological expressions are fundamental to human emotion recognition and regulation.

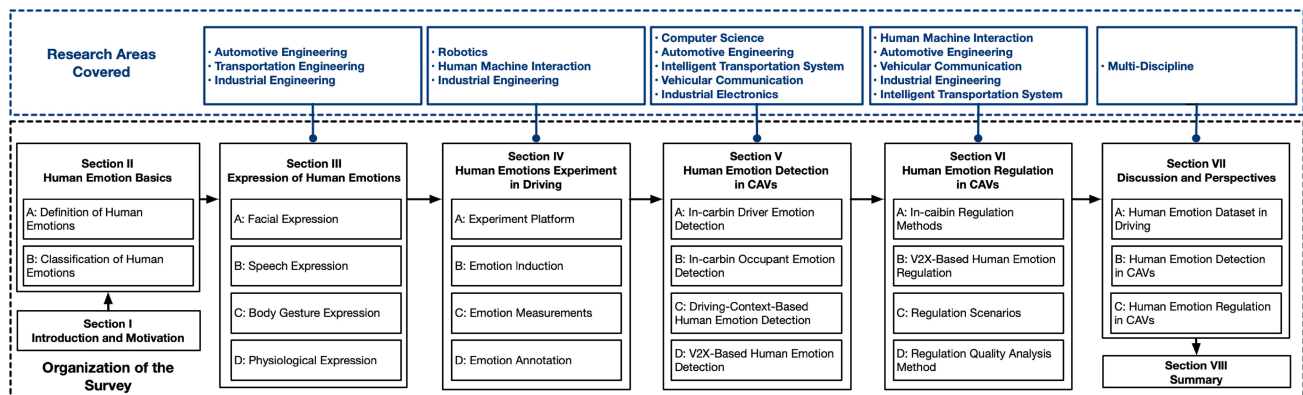
- Third, experimental methodologies for studying human emotions in driving are discussed, including experiment platform, emotion induction, emotion measurement, and emotion annotation.
- Then, how to detect human emotion effectively in CAVs is reviewed from four aspects: in-cabin driver emotion detection, in-cabin occupant emotion detection, driving-context-based driver emotion detection and V2X-based human emotion detection.
- Next, how to use the human-machine interface in CAVs to regulate the detected human emotion is discussed, consisting of in-cabin regulation methods, V2X-based human emotion regulation, regulation scenarios and regulation quality analysis method.
- Finally, the significant challenges and promising areas are highlighted.

## 2 Human Emotion Basics

The human emotion basics are a prerequisite for the study of human emotion in CAVs. In this section, the basic concepts of human emotions are discussed in terms of the definition and classification of emotions.

### 2.1 Definition of Human Emotion

Emotion research has been conducted for decades. However, researchers studying emotions still have not reached a broad consensus on how to define emotions precisely. Figure 3a presents the situation-attention-cognitive appraisal-emotional response sequence specified by the modal model of emotion in a highly abstracted and simplified form [17–21]. According to the cognitive theory [22],



**Fig. 2** Organization of the survey and related research areas

an emotional response begins with a cognitive appraisal of the personal significance of a situation, which in turn gives rise to an emotional response involving subjective experience, physiological change, and behavior response [17].

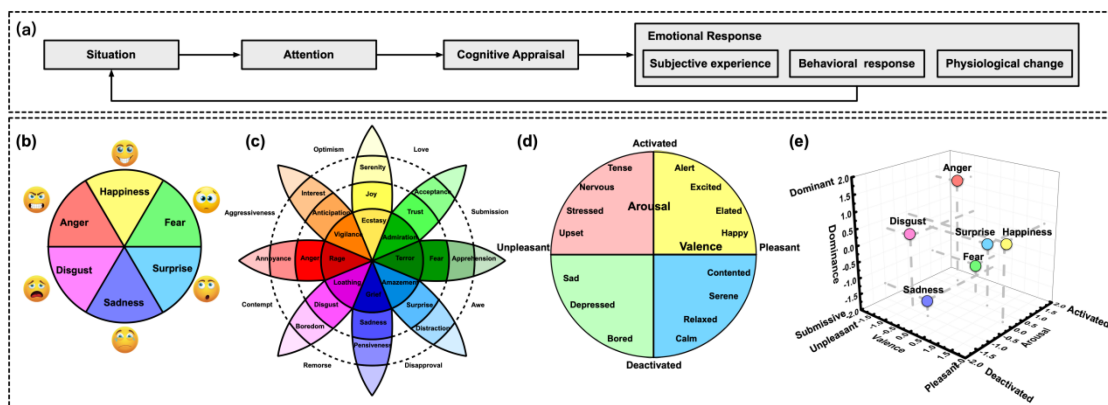
Cognitive appraisal refers to the personal interpretation of a situation or event. The appraisal process includes: (a) whether a situation or event threatens people’s well-being, (b) whether there are sufficient personal resources available for coping with the demand of the situation and (c) whether the strategy for dealing with the situation is effective [22]. Subjective experience shows that all humans have several basic universal emotions (e.g., anger, happiness, sadness) regardless of culture and ethnicity, and the way to experience these emotions is highly subjective [23]. Physiological change is a biological awakening or a physical reaction experience during an emotion. For example, scare usually leads to heart racing, short breathing, drying mouth, tensing muscles, and sweating palms [24]. Behavioral response leads to behavior that is often, but not always, expressive, goal-directed, and adaptive [25], for example, a smile indicating happiness and a frown indicating sadness. Additionally, emotional responses often change the situation that gives rise to the response in the first place. Figure 3(a) also depicts this aspect of emotion by showing the response looping back to (and modifying) the situation that gives rise to the emotion.

### 2.2 Classification of Human Emotion

Emotions can be classified into different categories according to discrete emotion theory or dimensional emotion theory. Discrete emotion theory assumes that there exists a small number of separate emotions, characterized by

coordinated response patterns in physiological, brain, and facial expressions. The current discrete emotion models mainly include the six emotion models (Fig. 3(b)) proposed by Ekman [18] and the emotion wheel model (Fig. 3(c)) proposed by Plutchik [19]. Among these two discrete models, the most acknowledged one is the basic emotion model. Ekman summarized the six basic emotions as happiness, sadness, anger, fear, surprise, and disgust. Other emotions were regarded as combinations of these basic emotions. In particular, a cross-cultural study conducted by Ekman supported the discrete emotion theory based on the six basic emotions [26].

Dimensional emotion theory proposed that emotional states can be accurately represented as a combination of several psychological dimensions (e.g., valence, arousal). The frequently used dimensional models include the circumplex model (Fig. 3(d)) and the 3D PAD emotion model (Fig. 3(e)). Most of these dimensional models contain valence and arousal. Valence refers to the degree of pleasure associated with emotions, whereas arousal refers to the intensity of the experienced emotions. Russell proposed the circumplex model in which emotions formed a circle defined by the dimensions of valence and arousal [27]. In the circumplex model, the subjective experience of pleasure and arousal is called core emotion [28]. Valence and arousal are useful as a basic taxonomy of emotional states [29], however, using only valence and arousal to project a high-dimensional emotional state onto a basic 2D space would result in information loss, for example, some emotions become indistinguishable (e.g., fear and anger), and some emotions are outside the space (e.g., surprise) [30]. Further, Mehrabian [21] extended the emotion model from 2D to 3D to develop a well-known pleasure-arousal-dominance emotion model.



**Fig. 3** Modal model and classification of human emotion. **a** is the emotion modal model [17] **b** and **c** are two discrete emotion models: **b** is the six basic emotions model [18], **c** is the emotion wheel model

[19]. (d) and (e) are two dimensional emotion models, **d** is the 2D circumplex model adapted from [20], **e** is the 3D PAD emotion model [21]

See Fig. 3(e) for illustration. The increased dimensional axis is named dominance. Ranging from submissive to dominant, dominance reflects the control ability of humans on emotions. In this model, anger and fear can be easily identified because anger is in the dominance axis, while fear is in the submissive axis.

- *Summary of human emotions basics:* Discrete emotions such as anger and sadness have been widely studied for a long time in the past, and are considered to be the primary extreme negative emotions that lead to traffic accidents. In the study of human emotions in driving, the widely used discrete emotion method can intuitively reflect emotions, but only includes a limited number of emotions, and it is difficult to clearly express mixed emotions. Dimensional emotion methods have the advantages of being practical and context-sensitive, but are less intuitive and require a more complex data labeling process. Therefore, both discrete and dimensional emotions should be discussed in CAVs' emotion research. Moreover, in previous studies, the main emotions studied in driving are the driver's negative emotions, and the positive emotions of drivers and occupants should also receive more attention in the research of CAVs.

### 3 Human Emotion Expressions

Human emotion expressions are fundamental to emotion recognition and regulation in driving. Emotional expressions are multimodal, involving the face, speech, body gesture, and physiological changes [31]. Therefore, human emotion expressions will be discussed from the above aspects in this section, separately.

### 3.1 Facial Expression

The association between facial expression and emotion is investigated by a series of studies in psychology [34]. Ekman's study shows that at least six basic emotions (anger, fear, disgust, happiness, sadness, and surprise) could be recognized cross-culturally by facial behavior [26]. Many of the relevant studies have quantified facial behavior using componential coding. In a further study, Ekman developed a facial motion coding system (FACS), the most comprehensive and widely used taxonomy in facial behavior coding [32]. FACS is an anatomically-based comprehensive measurement system that evaluates 44 different muscle movements (e.g., raising of the eyebrows, tightening of the lips). The action units (AUs) of FACS are suitable for studying human naturalistic facial behavior as the thousands of anatomically possible facial expressions can be described as a combination of 27 basic AUs and multiple AU descriptors. Table 1 illustrates the descriptor for each action unit and the AUs that can be observed in basic emotions [32]. Each AU corresponds to a specific facial movement, and different combinations of AUs represent different emotional expressions. Table 1 also presents the prototypical AUs observed in each basic emotion [33].

Most of the existing emotion studies on facial expression employed discrete emotion models, focusing on the expression of six basic emotions, due to their universal properties, and their apparent reference representations in human' emotional life [35]. There are also some studies analyzing the relationship between facial expressions and emotions based on dimensional emotion models. Russell's research shows that facial expression appears to indicate the valence of a person's emotional state reliably [36]. For example, Duchenne smiles [37] involving the wrinkling of muscles around the eyes are reported to associate

**Table 1** Action units (AU) descriptors [32] and prototypical AUs observed in each basic emotion [33]

AU	Descriptor	AU	Descriptor	AU	Descriptor	AU	Descriptor
1	Inner brow raiser	9	Nose wrinkler	16	Lower lip depressor	23	Lip tightener
2	Outer brow raiser	10	Upper lip raiser	17	Chin raiser	24	Lip pressor
4	Brow lowerer	11	Nasolabial fold deepener	18	Lip pucker	25	Lips part
5	Upper lid raiser	12	Lip corner puller	19	Tongue show	26	Jaw drop
6	Cheek raiser	13	Sharp lip puller	20	Lip stretcher	27	Mouth stretch
7	Lid tightener	14	Dimpler	21	Neck tightener	28	Lip suck
8	Lips toward each other	15	Lip corner depressor	22	Lip funneler	–	–
Basic Emotions		Anger	Sadness	Fear	Surprise	Disgust	Happiness
Prototypical AUs Observed		4, 7, 24	4, 15	1, 4, 20, 25	1, 2, 25, 26	9, 10, 17	12, 25

with positive emotional experiences. In contrast, negative emotion inductions are often associated with visible facial behavior in which the eyebrows are lowered and brought closer together [38].

### 3.2 Speech Expression

Speech delivers emotional information through linguistic and acoustic messages that reflect how the words are spoken [39]. The linguistic content of speech carries emotional information. Some of the information can be inferred directly from the surface features of the words, which have been summarized in emotional words dictionaries [40]. The rest of the information lies under the text surface and can only be understood when considering semantic contexts (e.g., discourse information). Although the association between linguistic content and emotion is language-dependent, multiple cross-lingual models have been developed for emotion detection [41, 42]. In addition to the linguistic messages to be expressed, the acoustic messages can also deliver rich information (such as age, gender, and hometown). Acoustic expressions are described by speech prosody (pitch, loudness, rhythm). A happy voice usually has a higher pitch, and a louder and faster speech rate, whereas a sad voice has a lower pitch and slower speech rate.

There are two opinions on acoustic expressions [43]. The first one is that acoustic expression that could distinguish different basic emotional states which are discrete. Early researchers such as Darwin stated that specific intonations corresponded to specific emotional states [43]. Banse and Scherer [44] studied the relationship between 14 induced emotions and 29 acoustic variables. The authors found that a combination of ten acoustic properties can distinguish discrete emotions. Cowie's research [45] on qualitative acoustic correlations of basic emotions shows that the basic emotion-related prosodic features extracted from audio signals

include pitch, energy, and speech rate. Table 2 illustrates the changes in various acoustic features during the occurrence of six basic emotions [45, 46]. The pitch, intensity, duration, and spectral are shown for each emotion, with changes indicated as either an increase, decrease, or no significant change.

The other opinion is the acoustic expressions are different in the continuous valence and arousal. The most consistent association reported in the literature is the relationship between arousal and acoustic pitch. Higher arousal levels are linked to higher-pitched vocal samples [47]. Bachorowski and Owen [48] suggested that acoustic pitch can be used to assess the level of emotional arousal currently experienced by individuals. It is much more challenging to find acoustic features sensitive to valence [47] because valence correlates strongly with textual features from the spoken word [49].

### 3.3 Body Gesture Expression

Many studies showed that body language plays a vital role in emotional communication, which conveys specific information about a person's emotional state [50]. Body language includes different types of non-verbal indicators such as facial expressions, body postures, gestures, eye movements, touch, and the use of personal space [51]. Besides the face, hands describe the most abundant information of body language [52]. For example, people using open gestures are considered with positive emotions [53]. Head positioning also reveals much information about the emotional state. People are prone to speak more if the listener encourages them by nodding [53]. The above knowledge indicates that it is necessary to consider various parts of the body at the same time to correctly interpret body language as an indicator of emotional state [50].

**Table 2** Acoustic features of basic emotions [45, 46]



















Acoustics	Anger	Sadness	Fear	Surprise	Disgust	Happiness
Pitch	Increase in mean, Median range, Variability	Below normal Mean F0, Range F0	Increase in mean F0, Range F0, Perturbation, Variability F0 movement	Wide range, Median normal or higher	Decrease in Mean, Range slightly wider	Increase in mean, Range, Variability
Intensity	Raised	Decreased	Normal	–	Decreased	Increased
Duration	High rate, Reduced rate	Slightly slow, Long pitch falls	Increased rate, Reduced rate	Tempo normal, Tempo restrained	Low rate	Increased rate, Slow tempo
Spectral	High midpoint for av spectrum for nonfric portions	Decrease in high-frequency energy	Increase in high-frequency energy	–	–	Increase in high-frequency energy

Studies have investigated the body language emotional coding system [54], but a broad consensus has not yet been reached on how to code body posture and movement. Table 3 provides a summary of the general body expressions associated with basic emotions, with detailed descriptions of body, hand, arm, leg, and other expressions [50, 55]. Based on the discrete emotional model, Gunes and Piccardi [56] showed that using naive representations from the upper body can match body postures into six emotional categories, namely anger, disgust, fear, happiness, sadness, and uncertainty. Castellano et al. [57] showed that dynamic representations of body postures could be categorized as anger, joy, happiness, or sadness. Based on skeletal geometric features, Saha et al. [58] identified gestures corresponding to five basic

human emotional states, including anger, fear, happiness, sadness, and relaxation.

Researchers also studied the emotional expression of body language based on dimensional emotion models. Glowinski et al. [59] showed that essential groups of emotion, associated with the four quadrants of the valence-arousal space, can be distinguished from the representation of trajectories of head and hands from the frontal and lateral view of the body. The compact representations were grouped into clusters that divided the inputs into four categories based on their positions in the dimensional space, namely high-positive, high-negative, low-positive or low-negative. Kosti et al. [60] used features extracted from the body and background to study the intensity values, arousal, and dominance.

**Table 3** The general body expression for the basic emotions [50, 55]

Emotion	Representative Visual Illustration			Associated Body Expression
	Front View	Side View	Rear View	
Anger				<p>Body: Body spread.</p> <p>Hands: Hands on hips or waist. Closed hands or clenched fists. Palm-down posture. Lift the right or left hand up.</p> <p>Finger point with right or left hand. Finger or hand shaky.</p> <p>Arms and legs: Arms crossing.</p>
Sadness				<p>Body: Body dropped. Shrunken body. Bowed shoulders.</p> <p>Body shifted. Trunk leaning forward. Self-touch (disbelief), body parts covered or arms around the body or shoulders.</p> <p>Body extended and hands over the head.</p> <p>Hands: Hands kept lower than their normal positions, hands closed or moving slowly. The face is covered with two hands. Two hands touching the head and moving slowly. One hand touching the neck. Hands closed together.</p> <p>Others: Head bent.</p>
Fear				<p>Body: Conservative body posture.</p> <p>Hands: Hands clenched.</p> <p>Arms and legs: Arms and legs crossing and moving. Arms clenched, elbows dragged inward, bouncy movements, legs wrapped around objects.</p> <p>Others: Noticeably high heart beat-rate (visible on the neck). Breath held.</p>
Surprise				<p>Body: Body shift or backing. Body abrupt backward movement.</p> <p>Hands: One hand or both of them moving toward the head. Moving one hand up. Both of the hands touching the head. One of the hands or both touching the face or mouth. Both of the hands over the head. One hand touching the face. Self-touch or both of the hands covering the cheeks or mouth.</p> <p>Others: Head shaking</p>
Disgust				<p>Body: Backing. Body shifted. Orientation changed or moving to aside.</p> <p>Hands: Hands covering the neck. One hand on the mouth. One hand up. Hands close to the body. Hands covering the head.</p>
Happiness				<p>Body: Body extended., Looking around.</p> <p>Arms and legs: Arms open. Arms move. Legs parallel. Legs open. Legs may be stretched apart.</p> <p>Others: Eye contact relaxed and lengthened. Feet pointing something or someone of interest.</p>

### 3.4 Physiological Expression

Although an individual may not clearly express his or her emotions through speech, gestures, or facial expressions, changes in physiological patterns are inevitable and detectable [61]. Emotional experience is accompanied by a series of physiological expressions that will, in turn, enhance the human emotional experience. The human physiological expression can be described from three aspects, including autonomic nervous system (ANS) response and central nervous system (CNS) response.

#### 3.4.1 ANS Response

The ANS is a general-purpose physiological system responsible for regulating peripheral function [62]. The system consists of sympathetic and parasympathetic branches, which are usually associated with activation and relaxation, respectively. When a person is positively or negatively excited, the sympathetic nerves of the ANS are activated [63]. This sympathetic activation increases the respiratory rate, raises heart rate (HR), reduces heart rate variability (HRV), and raises blood pressure [64].

Some emotional researchers proposed that each emotion (e.g., sadness, anger, fear) had a relatively certain and reliable autonomic response pattern [65]. Table 4 summarizes various studies investigating the autonomic specificity of basic emotions, which demonstrate that different emotions are associated with specific physiological responses, including galvanic skin response, heart rate, blood pressure, finger temperature, and heart rate variability [66, 67]. Other researchers indicated that it might be best to observe ANS responding from the perspective of the dimensional emotion model such as arousal [68]. Lang and his colleagues [69] found that skin conductance level (SCL) increases linearly with the rated arousal of emotional stimuli. A meta-analysis showed that blood pressure, cardiac output, HR, and skin conductance response duration respond to emotional valence [70].

#### 3.4.2 CNS Response

The physiological correlation of discrete emotions is more likely to be found in brain activities rather than in peripheral physiological responses [71]. Therefore, emotional representations in the CNS response have been investigated using a variety of frameworks [72]. A single-system emotion model assumes that one neurological system drives the experience and expression of all emotions, whereas multisystem models conceptualize emotions and their correlations into discrete bins [73]. Recent advances in functional neuroimaging have provided an opportunity to delineate a greater central mechanism supporting emotional processes [74].

Neuroimaging studies have shown that human emotions are controlled by a circuit, the phases of which include the orbitofrontal cortex, ventromedial prefrontal cortex, amygdala, hypothalamus, brainstem, cingulate cortex, thalamus, hippocampus, nucleus accumbens, insula, cortical sensory layers, etcetera. These are the critical brain regions for emotional production, emotional experience, and emotional regulation [73]. Activation and inactivation of different brain regions may indicate that they play different roles in emotional processing [75]. The application of activation likelihood estimation (ALE) revealed consistent, dissociable activation of distinct areas for happiness, sadness, anger, fear, and disgust [76]. Emotional specificity was confirmed by comparing emotional categories to each other. Kirby and Robinson [77] replicated the work of Vytal and Hamann [76] using the BrainMap database and ten times the number of studies. Although evidence of different brain region networks was found in certain emotions, there were also brain regions that were always activated in all emotions, suggesting the existence of a multisystem model [77].

- *Summary of human emotions expressions:* This section discusses the multimodal expression of human emotions from four aspects: facial expression, speech, body posture and physiological changes. Facial expressions and speech expressions are the best and most commonly used in-cabin methods to detect human emotions due to their non-invasive properties. Besides, for the drivers' emotion detection, there is a strong correlation between body gestures and driving behaviors. Considering steering wheel angle and gas pedal angle can reflect the drivers' body gestures, the relationship between body gestures and human emotions is illustrated. Moreover, physiological expressions are detailed in this section. Wearable physiological devices can be used to measure the physiological expressions (e.g., heart rate and skin conductance activity) in driving, and human brain activities are commonly used for the golden standard (ground truth) in emotion detection. With the development of sensing devices, CAVs will provide more detection methods for human emotion expression-based research in driving.

## 4 Human Emotion Experiment in Driving

Collecting rich and reliable data is an essential step in the study of human emotions while driving. Therefore, the purpose of this section is to discuss how to design experiments to collect valid data. Table 5 presents a summary of human emotion experiments in driving, including the experiment platforms, induction methods, emotion measurement methods, and emotion annotation methods.

**Table 4** Relationship between ANS responses and basic emotions [67]

ANS Response		Anger	Sadness	Fear	Surprise	Disgust	Happiness
<b>Cardiovascular</b>							
HR	Heart Rate	↑	↓	↑	↑	↓	↑
HRV	Heart Rate Variability	↓	↓	↓		–	↓
LF	Low Frequency Spectral HRV			(–)			–
LF/HF	Low Frequency/High Frequency Ratio					(↓)	
PWA	P-Wave Amplitude			(↑)			
TWA	T-Wave Amplitude	↓		(↓)		(↑)	
LVET	Left Ventricular Ejection Time	↓		↓		(↓)	(–)
HI	Heather Index	↑		(↑)			
PEP	Preejection Period	↓		↓		(↓)	(↑)
SV	Stroke Volume	↓↑		↓		↓	(–)
CO	Cardiac Output	↓↑		↑		↑	(–)
SBP	Systolic Blood Pressure	↑		↑		↑	↑
DBP	Diastolic Blood Pressure	↑		↑		↑	↑
MAP	Mean Arterial Pressure			↑		↑	↑
TPR	Total Peripheral Resistance	↑		↓		↓	(↑)
FPA	Finger Pulse Amplitude	↓	↓	↓		↓	↓↑
FPTT	Finger Pulse Transit Time	↓		↓		↓	↑
EPTT	Ear Pulse Transit Time			↓		↓	↑
FT	Finger Temperature	↓	↓	↓	↓↑	↓	↑
HT	Forehead Temperature	↓↑				(↓)	
<b>Electrodermal</b>							
SCR	Skin Conductance Response	↑		↑		↑	
nSRR	Nonspecific SCR Rate	↑	↓	↑		↑	↑
SCL	Skin Conductance Level	↑	↓	↑	(↑)	↑	↑–
<b>Respiratory</b>							
RR	Respiration Rate	↑	↑	↑	↓–	↑	↑
Ti	Inspiratory Time	(↓)		↓–		–	↓
Te	Expiratory Time	(↓)		↓		–	↓
Pi	Post-inspiratory Pause Time	(↑)		(↑)	(↑)		(↓)
Pe	Post-expiratory Pause Time						
Ti/Ttot	Inspiratory Duty Cycle			(↑)		–	
Vt	Tidal Volume	↓↑	↑	↓↑	↑–	(↓)	↓↑
Vi/Ti	Inspiratory Flow Rate or Inspiratory Drive					–	(↑)
Ros	Oscillatory Resistance	(↑)				–	(↑)
pCO2	End-tidal Carbon Dioxide Partial Pressure			↓			

Note: ANS responses were defined as the response direction reported by the majority of studies (unweighted), with at least three studies indicating the same response direction. Arrows indicate increased (↑), decreased (↓), or no change in activation from baseline (–), or both increases and decreases between studies (↑↓). Arrows in parentheses indicate tentative response direction, based on fewer than three studies. Sadness was defined as deactivating sadness (non-crying); Disgust was defined as mutilation disgust

## 4.1 Experiment Platform

As shown in the platform column of Table 5, the experiment platforms used in human emotion experiments in driving include both driving simulators and real vehicles for on-road driving. Driving simulators are widely used in automotive research, providing simple data collection, low-cost testing, driver safety, adaptability, and a fully controllable environment.

Consequently, driving simulators are often used in driver emotion research [78]. Limitations of the driving simulators compared to the real driving environment include limited physical, behavioral, and perceived fidelity, which may lead to unrealistic driving behavior, degraded subjects' involvement level in the study, and hazards not taken seriously by the driver, thus affecting the reliability of the research results [79]. To date, even the most advanced driving simulators are not able

**Table 5** Summary of driver emotion detection in CAVs

M	Ref.	Emotion	No. Sub	Platform	Induce Method	Measurement	Annotation	Data content	Algorithm	Detection result
Facial-expression-based	[82]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	12	Simulator	Situational procedures	FEA	Self-reports	55 image sequences	GLFCNN+SVM	Acc: FER 2013 84.41%, CK+ 95.05%, KDEF 98.57%, KMU-FED 98.64% Acc: 88%-happiness, 85%-sadness, 93%-anger Acc: 96.6%
	[83]	Happiness, Sadness, Anger	20	-	-	FEA	Annotators	200 images	OpenCV	Acc: 88%-happiness, 85%-sadness, 93%-anger Acc: 96.6%
	[84]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	25, male 20, female 5	On-Road	Situational procedures	FEA, Self-reported scale	Self-reports	15,523 images	FERDERnet with Xception backbone	Acc: 96.6%
	[85]	Anger, Neutral	10	Simulator	Situational procedures	FEA, Self-reported scale	Self-reports	20 videos	CNN+Softmax classifier	Acc: 88.25%
	[86]	Anger, Disgust, Happiness, Fear, Sadness, Surprise, Neutral, Valence, Arousal, Dominance	40	Simulator	Video-audio clips	FEA, Self-reported scale	Self-reported, Annotators	240 video sequences, each sequence 15 s	CNN-based CogEmoNet	Acc: DEFE+ 35.1%, CK+ 90.7%, MSE: DEFE+ 4.22, DEAP 3.654, CCC: DEFE+ 0.221, DEAP 0.181
	[87]	Neutral, Happiness, Anger, Valence, Arousal, Dominance	60, male 47, female 13	Simulator	Video-audio clips	FEA	Self-reports	164 video sequences each sequences 30 s	SVM/Xception	Acc: SVM 56.36%, Xception 89.01%
	[88]	Surprise, Neutral, Sadness, Fear, Happiness, Disgust, Anger, Contempt	-	Simulator	Situational procedures	FEA	Annotators	~10177 images	Multi-EmoNet	Acc: 83.1%
	[89]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	-	Simulator	Posed	FEA	Annotators	135000 images	CycleGAN style transferring neural network	Acc: 74.58%
	[90]	Anger, Neutral	15, male 8, female 7	Simulator	Situational procedures	FEA (NIR and thermal camera)	Experiment scenarios	58926 NIR images 58926 thermal images	VGG face-16 model	Acc: 99.96%

Table 5 (continued)

M	Ref.	Emotion	No. Sub	Platform	Induce Method	Measurement	Annotation	Data content	Algorithm	Detection result
	[91]	Happiness, Bother, Concentrated, Confusion	10	On-road	Situational procedures	FEA	Annotators	30 video segments of 30 fps, each 8-20 min	SVM	Acc: 37%
	[92]	Anger, Disgust, Neutral	12	Simulator	Directed facial action	FEA (NIR camera)	Acted	20 video segments of 25fps, each 1 min	Linear-SVM	Acc: 85%
Speech-based	[93]	Neutral, Positive, Frustrated, Anxious	28, male 22, female 6	Simulator	Situational procedures	Speech, Self-reported scale	Self-report	7562 samples of 186.24 min speech material	Random forest, SVM	Acc: Random forest 52.8%, SVM 77.4%
	[94]	Anger, Neutral	8, male 5, female 3	Simulator	Standardized imagery	Radio Frequency, Self-reported scale	Self-reports	100000 samples	KNN	Acc: 82.9%
	[95]	Negative, Neutral	68, male 35, female 33	On-road	Situational procedures	Speech	Experiment scenarios, Annotators	Speech samples in UTDrive	Gaussian mixture model +SVM	Acc: 79%
	[96]	Anger, Happiness, Sadness, Neutral	11, male 8, female 3	On-road	Situational procedures	Speech	Annotators	-	Multi layer perceptron, ANFIS, GenSoFNN	Acc: Multi-layer perceptron 63.6%, ANFIS 66.9%, GenSoFNN 60.1%
DB	[97]	Happiness, Neutral	11	On-road	Situational procedures	Driving behavior	Experiment scenarios, Annotators	60 min with 200Hz sampling frequency	Multi-layer perceptron	Acc: 76.14%
Physiology-based	[98]	Anxiety	31, male 16, female 15	Simulator	video clips	PPG, EDA, Pupil size, and EEG	Experiment scenarios, Self-reports	Each participant 60 sequences	Logistic regression classifier	Acc: 77.01%
	[99]	Anger, Neutral	20, male 16, female 4	Simulator	Situational procedures	EDA	Experiment scenarios	20 samples, each 1 min	SVM	Acc: 85%
	[100]	Tension, Relaxation	24, male 13, female 11	On-road	Situational procedures	EEG, EDA, Self-reported scale	Self-reports	-	Logistic regression, K-means	Kappa index: Logistic regression 54.55%, K-means 27.45%
	[101]	Fear, Anger, Boredom	41, male 36, female 5	Simulator	Situational procedures	EDA, ECG, Respiration, EMG, Finger Pressure	Experiment scenarios, Self-reports	116 records, each record 12 features	KNN, Marquardt Backpropagation, Resilient Back-propagation	Acc: KNN 65.33%, MBP 73.26%, RBP 82.6%

Table 5 (continued)

M	Ref.	Emotion	No. Sub	Platform	Induce Method	Measurement	Annotation	Data content	Algorithm	Detection result
Multimodal-based	[102]	Anger, Neutral	30, males 30	On-Road	Situational procedures	EEG, BVP, HR, EDA, Skin Temperature, Respiration, Self-reported scale, Driving behavior	Self-reports	906 angry samples, 713 neutral samples	Least square support vector machine	Acc: 82.2%
	[103]	Joy, Anger, Sadness, Fear	30, male 18, female 12	On-Road Simulator	Pictures, Music, Film clips, Situational procedures	Self-reported scale, Driving behavior	Self-reports	481 groups of the actual and virtual driving data	Fuzzy Comprehensive Evaluation	Acc: Joy 82%, Anger 88%, Fear 72%, Sadness 75%
	[104]	Joy, Annoyance, Insecurity	36	Simulator	Situational procedures	FEA, Speech, Self-reported scale	Annotators	288 interactions data and agent 180 between driver and co-driver 108 between driver and co-driver	Neural transfer Learning	F1: 76%
	[105]	Neutral, Boredom, Frustration	10, male 7, female 3	Simulator	Situational procedures	PPG, EMG, IMU	Experiment scenarios	7200 data points, 3 min windows with 40 Hz	SVM	Acc: 99.52%
	[106]	Anger, Neutral	30, male 30	On-road	Situational procedures	EEG, BVP, HR, EDA, Skin Temperature, Respiration, Self-reported scale, Driving behavior	Self-reports	510 angry samples, 120 neutral samples	Linear discriminant model	Acc: 85.84%
	[107]	Frustration	30, male 20, female 10	On-road	Situational procedures	FEA, EDA, Self-reported scale, Driving behavior	Annotators	–	Bayesian network	Acc: 80%

to imitate highly complex and realistic driving situations [78]. Motion sickness is a common side-effect of subjects in simulator studies, leading to blurred vision, migraine, epilepsy, and dizziness, negatively affecting the usability of simulators [79].

Although many approaches have been conducted in the laboratory environment, a controlled environment still has a stronger influence on subjects' emotions and behavior, which may limit the reliability of the generated results. Therefore, many researchers conduct on-road testing at the advanced stage of research [80]. Currently, there is no framework for on-road research for human emotion investigations. Generally, when deciding on a driving route, subjects are often required to travel along a route familiar to them, or pre-planned routes to cover different driving situations and road types. Previous automotive research suggests combining three different road types for planning road routes or comparing them: rural, urban, and major [81]. However, a significant concern for on-road research is that it is unsafe for an emotional driver to drive on real roads that may cause traffic accidents.

## 4.2 Emotion Induction

The high quality of emotion elicitation stimuli is vital for driver emotion studies [102]. The purpose of emotion-inducing methods is to induce the target emotions of subjects. Generally, choosing the appropriate elicitation scene or stimulus depends on the target emotions and the experimental situation. As shown in the induce method column of Table 5, various emotion induction techniques have been used in human emotion experiments in driving, including emotional video clips, traffic scenarios, music, and other methods. Siedlecka and Denson [108] have classified emotion induction techniques into five broad categories: situational procedures, visual stimuli, auditory stimuli, standardized imagery, and autobiographical recall. Emotional behavior can also be used as an emotional stimulus.

Situational procedures involve creating a driving situation and putting subjects directly into the situation that elicits the target emotion, and is generally achieved by controlling the traffic, environment, accidents and personal interaction (e.g., catch a plane during the time of road congestion) [90, 102]. Visual stimuli can be film clips (videos) or static images with emotional content selected to elicit target emotions [103]. Auditory stimuli activate emotions via specific types of auditory input (e.g., music and the voice of the smart agent in the car) [103]. Standardized imagery involves subjects creating vivid mental representations and imagining themselves in a novel event designed to evoke strong emotion [94]. Standardized imagery can consist of reading vignettes, often with guidance from the experimenter. Autobiographical recall involves summoning personal emotional memories to reactivate emotions from the original emotional experience. Emotional behavior as emotional stimuli

involves participants posing specific expressions or behaviors (e.g., posing a smiling expression to induce happiness) to elicit target emotions [92]. Studies have shown that visual stimuli usually achieve better results in emotional induction among the basic emotions [108]. Besides, as shown in the induce method column of Table 5, situational procedures are mainly used for emotion induction in driving scenarios.

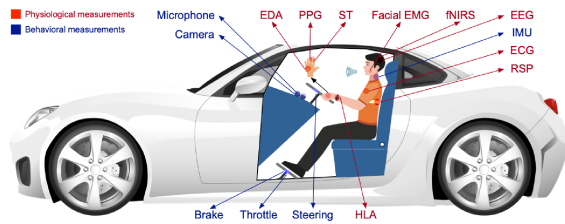
## 4.3 Emotion Measurement

Using appropriate methods to accurately measure various changes caused by emotions is essential for emotion detection. The measurement column of Table 5 lists various measurements that have been used in human emotion driving research. These measurements can be broadly classified into three categories: physiological measurements, behavioral measurements, and self-reported scales. Physiological measurements refer to the recording and analysis of biological signals from the human body. Behavioral measurements refer to the observation and analysis of human behavior in response to emotional stimuli. Self-reported scales refer to questionnaires or surveys that ask participants to rate their own emotional state in response to stimuli.

Since emotional features are closely related to the drivers' physiological characteristics, physiological measurements such as electroencephalography (EEG), functional near-infrared spectroscopy (fNIRS), electrocardiography (ECG), electrodermal activity (EDA), and respiration (RESP) used in driver physiology can be used to detect driver emotions [99–101]. There are also some authors who use self-reported scales to measure driver emotions. Frequently used self-reported scales include the positive and negative affect schedule (PANAS) [109], the self-assessment manikin (SAM) [110], and the differential emotions scale (DES) [111].

In addition, researchers have used behavioral measurements to infer driver emotions. To recognize emotions more accurately through facial behavior during driving, Gao et al. [92] used the method of near-infrared (NIR) facial expression recognition. Emotional acoustic recognition based on dialogue between the driver and the in-vehicle information system was also used to measure emotions [93, 95, 96]. Also, to identify drivers' emotions through body posture movement, Raja et al. [94] proposed to use radio frequency (RF) to identify driver emotions. Some authors suggest collecting information about the driver, vehicle state, and surrounding changes to infer the driver state. Nor and Wahab [97] classified the drivers' happiness and unhappiness emotion using a multilayer perceptron neural network based on the pressure characteristics of the gas and brake pedals.

Comprehensively considering emotion measurement from different aspects is vital for emotion understanding. In order to obtain more reliable results, different measurement



**Fig. 4** Position of the in-cabin sensors used in driver emotion detection

tools are usually combined for research [102–107]. These researchers used physiological measurements, behavioral measurements, and self-reported measurements to identify drivers' emotional states. For example, Wan et al. [102] reported the detection of anger using features of HR, blood volume pulse (BVP), skin temperature (ST), EDA, RESP, self-reported scale and driving behavior of drivers. Lee et al. [105] reported the detection of three basic emotions (neutral, boredom, frustration) using features of photoplethysmography (PPG), electromyography (EMG), as well as the IMU. The position of the in-cabin sensors used in driver emotion detection is illustrated in Fig. 4. The application of these methods has made massive progress in effectively monitoring and detecting driver emotions in real time.

Comparatively, physiological measurements are more objective and can be measured continuously; however, this measurement is highly intrusive and may have an impact on the measurement result. Although behavioral measurement is rarely intrusive and can be measured by non-contact sensors, its performance level could decline because people can try to fake or disguise their emotions. Self-reported scales measure the subjective experience of participants when applied correctly, and provide the opportunity to gain valuable insights about the experienced emotions, but such measurements cannot take place during the study without interruption. The methods require a relatively long period to perform, potentially resulting in unreliable data. In general, in order to obtain more accurate recognition results, an appropriate measurement method needs to be selected according to the experimental environment and purpose.

#### 4.4 Emotion Annotation

To effectively detect human emotions in CAVs, it is important to obtain reliable emotional annotations which can be used as the standard for training and evaluation of the models. As shown in the annotation column of Table 5, there are three annotation methods: self-reports, experimental scenarios, and annotators. The first approach involves employing emotional self-reports [87, 103], which requires subjects to be able to express and quantify how they feel at a particular moment. The second approach involves using

different experimental scenarios to label the emotional experiences [90, 99]. For example, in the case of on-road studies, researchers control different driving traffic to trigger specific emotional states, and in the case of driving simulator studies, researchers have several opportunities to modify experimental situations and induce specific emotional states. The third method involves the use of an external annotator, which can identify certain emotional states based on different signals (e.g., voice, facial expressions) of subjects [96, 97].

In comparison, self-reports require the cognitive recourse of subjects, they are subjective, and they may reflect strong biases (e.g., inconsistencies in evaluation criteria, false memories, and desire to impress the experimenter). Although the experiment scenarios reduce the burden on the subjects, it makes some strong assumptions that may not be widely generalized to all subjects and road conditions [15]. Annotators are very time-consuming and labor-intensive and require the use of experienced and well-trained observers, which may be challenging to find and will be expensive, especially at scale [15]. Besides, some studies have also explored combining multiple methods to overcome the weaknesses of any one method [112]. On the whole, no one method is perfect, so researchers need to choose a proper annotation approach according to the experimental scenario and purpose.

- Summary of human emotion experiment in driving:* This section discusses the platform, induction, measurement, and annotation in human emotion experiments. The collection of emotional data during driving mainly depends on the experimental platform. The data collection based on the driving simulator can reduce the driving risk caused by emotion during the experiment. Data collection based on on-road driving can collect more accurate responses from participants in real scenarios. At present, most studies choose to carry out driving simulator experiments in the early stage of emotion research and choose to carry out on-road driving experiments in the later stage or verification stage of emotion research. For the induction, measurement, and labeling of emotions, the choice should also be made according to the experimental platform. On the driving simulator, the experimental emotion measurement tools are rich, the induction method is controllable, and the labeling is easy, but there will be problems that the intensity of the collected emotional data is not high and the characteristics are not obvious. On the on-road driving, the emotions experienced by participants are more real, and the intensity of the collected emotional data is relatively high, but the emotional measurement tools of the experiment are easily limited (measurement equipment is easily disturbed in an uncontrollable environment), and labeling also requires a lot of time and human resources. Therefore, in the human

emotion experiment in driving, it is necessary to combine the research stage and the purpose of the experiment, choose the appropriate experimental platform, and further choose the measurement method, induction method, and labeling tools for emotion research.

## 5 Human Emotion Detection in CAVs

Accurate, stable, and efficient detection of the emotional states of drivers and occupants is a prerequisite for affective human-vehicle interactions. Therefore, the purpose of this section is to discuss how CAVs use emotional expressions, driving context, and road information to recognize the emotional states of drivers and occupants.

### 5.1 In-cabin Driver Emotion Detection

Driver emotion detection algorithms are developed based on selected classification methods and measurement signals. The algorithms are usually implemented using supervised machine learning technologies. The algorithm column of Table 5 not only lists the different measurements used in driver emotion driving research, but also summarizes the main detection algorithms and corresponding recognition accuracy reported in previous studies. The detection algorithms used in these studies can be broadly classified into three categories: facial-expression-based, speech-based, driving-behavior-based, physiology-based, and multimodal-based.

#### 5.1.1 Facial-Expression-Based

Facial expression is an effective, common-used, and non-invasive factor for emotion detection. There are two main detection tasks for emotion detection: discrete emotion detection (classification task for basic emotions) and dimensional emotion detection (regression task for valence, arousal and dominance). As shown in Table 5, most of the facial-expression-based studies choose the discrete emotion (six basic emotions and neutral) as the detection targets, and the deep neural network, such as CNN-based models, are commonly used in facial classification, such as GLFCNN [82], Xception [86, 87], and basic CNN [85, 90]. Equations (1) and (2) show the most commonly used loss function (multi-class cross-entropy loss and binary cross-entropy loss) for facial-expression-based emotion detection [82, 85, 86, 88]. As for dimensional emotion detection, some researchers choose the regression methods for emotion detection. The CogEmoNet was proposed to solve the regression task for

dimensional emotion recognition, and the Mean Squared Error (MSE) and Consistency Correlation Coefficient (CCC) loss functions were used [86]. The loss function of MSE and CCC are shown in Eqs. (3) and (4):

$$Loss_{\text{multi-class}} = - \sum_{i=1}^n y_i \log \hat{y}_i \quad (1)$$

$$Loss_{\text{binary}} = - \frac{1}{n} \sum_x [y \ln a + (1 - y) \ln (1 - a)] \quad (2)$$

$$Loss_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^M (I_t^i)^2 \quad (3)$$

$$Loss_{\text{CCC}} = 1 - \frac{2S_c}{S^2 + \hat{S}^2 + (\bar{y} - \bar{\hat{y}})^2} \quad (4)$$

Besides, some machine learning methods are used in facial expression classification, such as SVM [87, 91, 92] [82]. The loss function of SVM used in the detection procedure is hinge loss. According to Table 5, the results of CNN-based methods are generally better than those using traditional machine learning methods in driver emotion studies. CNN is a type of deep learning algorithm that can automatically learn and extract features from input data, making them well-suited for facial-expression-based emotion recognition tasks. In contrast, traditional machine learning methods require manual feature extraction and selection, which can be time-consuming and may not capture all relevant features. Using CNN-based methods can improve the accuracy of driver emotion recognition tasks and may be a promising approach for future research in this field.

#### 5.1.2 Speech-Based and Driving-Behavior-Based

The speech data is also a practical and non-invasive method for driver emotion detection. Current speech-based studies are usually used to detect drivers' road rage and negative emotions. The classification methods used are mainly traditional classification methods, such as KNN [94], GMM [95], SVM, and Random forest [93]. The commonly used loss functions for GMM, SVM, and random forest are log-likelihood loss, Hinger loss, and MSE loss, respectively. The log-likelihood loss is shown in Eq. (5). Besides, some researchers used neural networks to solve classification tasks. Kamaruddin et al. [96] used MLP to perform 4-class emotion classification based on audio information. For driving behavior data of drivers, Nor and Wahab used MLP to classify the happiness and neutral emotions [97].

$$Loss_{\log\text{-likelihood}} = \sum_{j=1}^N \log \sum_{k=1}^K z_k p(x_j | z_k, \theta_k) \quad (5)$$

### 5.1.3 Physiology-Based

The physiological signal is closely related to the drivers' emotional features. Therefore, some studies attend to use physiological data to detect drivers' emotions. The Logistic regression [98, 100], SVM [99], and backpropagation methods [101] are commonly used in the study. However, the number of studies addressing such classification tasks is fewer than other approaches. Since the physiological signal devices, such as ECG, and EEG devices, are not non-invasive. Meanwhile, high-quality physiological data is difficult to obtain compared to facial expressions and audio signals.

### 5.1.4 Multimodal-Based

The uni-modal approaches such as facial expression, audio, driving behaviors, and physiological data have their limits. Generally, the reliability of the uni-modal approaches is not good enough, especially for practical use. Therefore, some researchers choose multimodal signals to address the classification tasks. The most commonly used combinations, as shown in Table 5, are physiological data and driving behaviors data [102, 105–107], speech signals and driving behaviors data [103], and facial expression and speech signals [104, 107]. These combinations are often used in multimodal approaches, which combine multiple sources of data to improve the accuracy of emotion detection. The traditional machine learning methods, such as SVM [105], linear discriminant [106], least square support vector machine [102], and bayesian network [107] are commonly used in solving this task and the loss functions are hinge loss, and log loss function, respectively. These methods have been applied to different modalities and have achieved varying levels of accuracy in driver emotion detection. Transfer learning algorithms have also been successfully implemented for driver emotion detection, as shown in one study using a transfer learning approach to classify emotions based on facial expressions and speech signals [104]. Overall, the results of multimodal approaches are generally better than those of unimodal approaches, suggesting that combining different sources of data can improve the accuracy of emotion detection in driving. However, the choice of methods and combinations may depend on the specific context and available data.

## 5.2 In-cabin Occupant Emotion Detection

With the development of CAVs, emotion recognition for vehicle occupants gradually gains more attention. Since the

current CAVs have not reached the level of fully autonomous driving, only a few numbers of researches on occupants' emotion detection have been conducted. However, the models of general emotion detection might be able to be used for occupant emotion detection since there is no driving task involved for occupants. Table 6 presents a summary of the state-of-the-art emotion detection algorithms for occupants in vehicles in the past five years. Compared to the detection models used for drivers' emotion detection, the detection models used for general emotion detection are more up-to-date and the accuracy is generally higher. In Table 6, the occupants' emotion detection methods are categorized into four parts corresponding to the driver emotion detection, including facial-expression-based, speech-based, physiology-based, and multimodal-based. Then, each column in the table is discussed by category in the following subsections.

### 5.2.1 Facial-Expression-Based

In Table 6, facial expressions are one of the modalities used to detect the emotions of vehicle occupants. Some studies are summarized in the facial-expression-based part and the datasets used in these studies are the CK+ dataset, Japanese Female Facial Expression (JAFPE) Dataset, and CREMA-D dataset, which contain the facial expression data and labeled emotion annotations. The facial expressions in CK+, JAFPE, AFEW, and CREMA-D are acted by actors/actresses [113, 114, 116]. In terms of algorithms, Table 6 shows that the detection accuracies are generally higher when the expression datasets are acted compared to the dataset of spontaneous expressions. The reason is that acted facial expressions are more exaggerated and easier to recognize than spontaneous expressions, which tend to be more subtle and variable. However, spontaneous expressions are more reflective of real-world emotions and are therefore more relevant to emotion detection in naturalistic settings such as driving. Meanwhile, the algorithms shown in Table 6 are mainly CNN- and RNN-based models. The MSE and cross-entropy loss are commonly used as the loss function [117] in facial expression-based emotion detection. The loss functions are shown in Eqs. (1), (2), (3), and (4). Dimensional emotional recognition is not studied in occupants' emotion detection and the reason might be due to the limitation of the public dataset that contains the facial expression and arousal and valence scores.

### 5.2.2 Speech-Based

Table 6 shows that some studies have employed speech signals as the approach to detect the emotions of vehicle occupants. Speech recognition systems can analyze the tone, pitch, and prosody of the driver's voice to detect emotions such as frustration, excitement, and anger. They are summarized in the speech-based part and the major datasets used in

speech-based studies are RAVDESS, IEMOCAP, eINTERFACE, and RECOLA database, which contain the emotional speech segments and the labeled emotion annotation. The speech records in the RAVDESS dataset, IEMOCAP dataset, and eINTERFACE dataset are acted by actors/actresses [119, 122] and the speech records are spontaneous in RECOLA database [123]. The algorithms shown in Table 6 are mainly CNN-based models and RNN-based models, such as Bi-directional LSTM architecture [119], Deep C-RNN [120], etc. The MSE [123] and cross-entropy [121] loss are commonly used as the loss function in speech-based detection.

### 5.2.3 Physiology-Based

As shown in Table 6, physiological signals such as EEG are used to detect the emotions of vehicle occupants. Sensors can be installed on the vehicle to measure these physiological signals, which can then be analyzed using machine learning algorithms to detect emotions. Previous studies are summarized in the physiological-based section. The major datasets used in physiology-based studies are SEED dataset [124, 125, 128], and DEAP dataset [129], which contain the emotional physiological signals, the labeled emotion and the dimensional emotion annotation. Most of the datasets choose EEG signals as the emotion-related factor and eye movement is also contained in some datasets, such as SEED dataset [126, 127]. Besides, the algorithms used in physiology-based detection are mainly CNN-based models and transfer learning models.

### 5.2.4 Multimodal-Based

Some studies also choose multimodal signals as the approach to detect the emotions of occupants. Table 6 summarizes studies that have employed multimodal approaches to emotion detection for vehicle occupants. These approaches combine multiple modalities, such as facial expressions, speech signals, and physiological signals to improve the accuracy of emotion detection. The major datasets used in these studies are the IEMOCAP and MOSI datasets, which contain speech signals, facial expressions, and labeled emotion annotations. Deep learning algorithms such as LSTM-based models have been employed in some studies to analyze these multimodal datasets and achieve high accuracy in emotion detection [131–133]. Typically, the number of multimodal detection studies is less than the uni-modal study. The reason might be due to the limitation of the public datasets that contain facial expressions and audio signals. The major loss function used in the multimodal approach is MSE [130] and cross-entropy [132, 133].

## 5.3 Driving-Context-Based Human Emotion Detection

Since human emotions are mainly triggered by the driving context or the events that happened during their driving, such as the weather, the road type, and the traffic conditions. The driving context is also an important factor for drivers' emotion recognition. Although the studies towards such areas are limited because traditional vehicles lack of capacity to perceive the condition outside the vehicles, with the development of CAVs and the increase of the perception of automated vehicles, the combination of driving context and in-cabin collected signals will become possible and the accuracy of detection may be further improved.

Table 7 summarizes the current studies on the driving-context-based approach for drivers' emotion detection. These studies focus on analyzing the influence of different driving contexts such as traffic conditions and road types on drivers' emotional states. Bustos et al. [134] and Bitkina et al. [138] are examples of researchers that have investigated the influence of traffic conditions and road types on drivers' stress levels. These studies have found that certain driving contexts such as heavy traffic or narrow roads can lead to increased stress levels in drivers. Logistic regression was used to detect the stress levels and the highest accuracy reached 82.80%. Dobbins et al. [137] used the combination of the physiological data and driving context to analyze the stress level, and correlation analysis was performed in their discussion. Bethge et al. [136] proposed a new method to use the traffic condition, weather, and vehicles' dynamic information to detect drivers' emotions, and VEmotion model was used to detect the emotions and the accuracy reached 71.00%. Liu et al. [135] used an RNN model to detect the drivers' emotions by using the surrounding objects and vehicles' dynamic information. The accuracy of the detection reaches 71%. The number of studies in driving-context-based emotion detection is limited according to Table 7, and further study in this area is needed with the development of CAVs.

## 5.4 V2X-Based Human Emotion Detection

Integrating V2X technology into driver emotion detection, a driver can recognize his neighbors' emotional states in real-time by exchanging intelligent on-board data with other intelligent devices in ITS. For example, each CAV can share its emotional awareness information by interacting with other vehicles [139]. Furthermore, a local vehicle's on-board learning data can be combined and processed by the centralized cloud so that to extend its intelligence to the global context [140]. Other emerging technologies are integrated with

**Table 6** A brief summary of human emotion detection

M	Ref.	Emotion	Condition	Measurement	Annotation	Algorithm	Dataset	Detection result
Facial-expression-based	[113]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	Posed	FEA	Annotators	Dual integrated CNN	CK+	98.54%
	[114]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	Posed	FEA	Annotators	CNN- LSTM based Neural Network	CREMA-D	78.52%
	[115]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	Posed	FEA	Annotators	Deep Convolutional Neural Networks	JAFPE	95.23%
	[116]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	Posed	FEA	Annotators	CNN+LMED+DL-a+LSTM	AFEW	61.87%
	[117]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	Posed	FEA	Annotators	Convolution-Recurrent Neural Network	JAFPE	94.91%
	[118]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	Posed	FEA	Annotators, Self-reported	FER CNNs	CK+	96.57%
Speech-based	[119]	Calm, Happiness, Sadness, Anger, Fearful	Posed	Speech	Annotators	Bi-directional LSTM architecture and deep belief networks	RAVDESS	82.00%
	[120]	Calm, Happiness, Sadness, Anger, Fearful	Posed	Speech	Annotators	Deep C-RNN	RAVDESS	80.00%
	[121]	Anger, excitement, Neutral, Sadness	Posed	Speech	Annotators	self-attentional CNN-BLSTM model	IEMOCAP	81.60%
	[116]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	Posed	Speech	Annotators	Brain emotional learning model	CASIA, SAVEE, FAU	90.28%, 76.40%, 71.05%

**Table 6** (continued)

M	Ref.	Emotion	Condition	Measurement	Annotation	Algorithm	Dataset	Detection result
Physiology-based	[122]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	Posed	Speech	Self-reported, Annotators	Transfer learning	RML, eINTERFACE	49.70%, 34.30%
	[123]	Arousal, Valence	Spontaneous	Speech	Self-reported	2-layer Long Short-Term Memory	RECOLA	Arousal 78.70% Valence 44.00%
	[124]	Neutral, Sadness, Fear, Happiness	Spontaneous	EEG	Annotators	Coupled Projection Transfer Metric Learning	SEED_IV	82.16%
	[125]	Positive, Neutral, Negative	Spontaneous	EEG, Eye move	Self-reported	Deep neural Network	SEED (Chinese, Ger, Fra)	Chinese—86.53%, German—70.87%, French—67.52%
	[126]	Positive, Neutral, Negative	Spontaneous	EEG, Eye move	Self-reported	Multimodal domain adaptive variational autoencoder	SEED	89.64%, 73.82%
	[127]	Happiness, Sadness, Disgust, Neutral, Fear	Spontaneous	EEG, Eye move	Self-reported	Semisupervised sparse low-rank regression	SEED_V	79.02%
	[128]	Neutral, Sadness, Fear, Happiness	Spontaneous	EEG	Annotators	sJSFE	SEED-IV	82.45%
	[129]	Arousal, Valence	Spontaneous	EEG	Self-reported	SparseDGCNN	DEAP, SEED, DREAMER, CMEED	Averagely improved accuracy by 8.88%
Multimodal-based	[130]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	Posed	FEA, Speech	Annotators	Semi-supervised Multimodal Interaction Network	IEMOCAP	87.43%
	[131]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	Posed	FEA, Speech	Annotators	Reinforcement learning Dueling deep-Q-network	IEMOCAP	59.72%
	[132]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral	Posed	FEA, Speech	Annotators	Multi-layer LSTM	MOSI	77.30%
	[133]	Anger, Excitement, Neutral, Sadness	Posed	FEA, Speech	Annotators	LSTM	IEMOCAP	71.04%

**Table 7** Summary of driving-context-based human emotion detection

Ref.	Emotion	No. Sub	Platform	Driving context	Induce Method	Measurement	Annotation	Algorithm	Detection result
[134]	Stress	9	On-Road	Z zone, highway and city	Situational procedures	Front-view Video	Annotators	CNN	72%
[135]	Anger, Disgust, Happiness, Sadness, Surprise, Fear, Neutral, Valence	9, male 5, female 4	On-Road	Small vehicles, trucks, pedestrians, cyclists, location and CAN info from vehicles	Situational procedures	Front-view Video, CAN info from vehicles	Annotators	RNN	71%
[136]	Angry, Disgust, Happiness, Surprise, Calm, Fear Sadness, Neutral	12, male 8, female 4	On-Road	Weather, traffic flow, road type, vehicle state	Situational procedures	Weather, traffic flow, road type, vehicle state and personal facial expression, audio	Annotators	VEmotion	72%
[137]	Stress	8, male 2, female 6	On-Road	Idling, journey impedance, high urban workload, low urban workload, high non-urban workload, low non-urban workload	Situational procedures	Vehicle speed, location, photographs, ECG, PPG	Self reports	Data analysis	–
[138]	Stress	1 male	On-Road	Traffic conditions (average traffic speed) road types(city, highway)	Situational procedures	Front-view Video and EDA, PPG	–	LR	82.8% 80.3%

V2X to improve the efficiency of drivers' emotion detection, such as machine learning [141] and edge computing [142].

In Ref. [139], a real-time anomaly data detection method is proposed for safe driving based on the driver's emotional state and the cooperative vehicular networks. Zhang et al. [140] proposed a safety-oriented vehicular controller area network (SOVCAN) based on driver emotion detection. The architecture of SOVCAN consists of a data acquisition layer, an intra-net layer, and an inter-net layer. At the data acquisition layer, the driver's physiological and psychological data are collected by the in-car mobile. At the intra-net layer, the data is transmitted to the on-board device for emotion perception. The processed data is then transferred to the cloud through vehicle-to-cloud (V2C), vehicle-to-infrastructure (V2I), and vehicle-to-vehicle (V2V) communications for preservation and remote control. Vogel et al. [141] introduced a multi-layer cognitive architecture-emotion-aware vehicle assistant (EVA) for the future design of emotion-aware CAVs. With the emergence of edge computing, Raja et al. [142] designed a driver assistant system RExpress for emotion detection based on wireless edges.

- *Summary of human emotion detection in CAVs:* This section discusses methods for recognizing human emotions in CAVs. In the emotion recognition subsections for the driver and occupant in the cabin, the existing emotion recognition models are illustrated based on facial expressions, speech, behavior, physiological changes and multimodal. In addition, as recently focused areas, human emotion recognition based on driving environment information and V2X is also introduced. On the whole, considering that the CAV is an intelligent mobile terminal with information perception inside and outside the vehicle and information communication capabilities, the human emotion detection method in CAVs can use the fusion of environmental information inside and outside the vehicle to achieve accurate, stable and efficient recognition.

## 6 Human Emotion Regulation in CAVs

After detecting human emotions, the CAV needs to use the human-machine interface to regulate driver emotions for affective human-vehicle interactions. Thus, different

displays have been employed to regulate drivers' emotions in the driving context, as shown in Table 8. The display is an essential component of driver's emotion regulation, including visual displays such as screens and AR head-up displays, tactile displays such as the vibration of seats, steering wheels, and seat belts, olfactory displays such as lavender and lemon scents, and auditory displays such as voice output. Besides, this section introduces road-related regulation methods, regulation scenarios, and regulation quality analysis methods.

## 6.1 In-cabin Human Emotion Regulation

### 6.1.1 Visual Regulation

Vision is the dominant sense during driving, which is the reason for the majority of emotion regulation to be visual. For a low automated level, the use of visual stimuli for driver emotion regulation may cause potential distractions. However, a relaxation of visual regulation requirements is possible when the vehicle runs at a high level of automation. Table 8 describes current studies on the visual regulation of driver's emotions including ambient light, visual intervention, state feedback, and visual relaxation techniques.

Ambient light refers to the method employing ambient illumination in the cockpit to implement the driver's emotion regulation. Studies have shown that ambient light can have alarming or distracting qualities for drivers, but it may also have a calming effect, which may depend on brightness, position, and personal familiarity with the function [145, 149, 170]. Visual intervention is distractive by design, therefore, their application while driving requires to be carefully planned. It is also feasible to blatantly intervene when the possibly dangerous emotional state is likely to occur, such as providing a simple notification, telling the driver to take a break, and distracting the driver from the source of negative emotions, as well as presenting a positive notification to decrease drivers' negative emotions [143, 146, 147]. State feedback refers to the visual display that allows drivers to understand their current emotional states clearly. Researchers' studies found that direct feedback on the detected drivers' states had little value for emotional regulation because visual state feedback could amplify the driver's negative emotional states, which was unacceptable to the driver and needed to be avoided examined the visual state feedback of drivers' negative emotions [143, 147, 170]. Besides, participants preferred to receive only safety-critical notifications of the driver's state [144]. Relaxation techniques are approaches to impact drivers' emotional states through dynamic behavior changes. Studies have revealed that this relaxation technique led to a decrease in arousal levels and relaxation techniques may become a valuable routine to increase the comfort of CAVs [148].

### 6.1.2 Tactile Regulation

In the driving context, the tactile display can be presented in multiple ways (e.g., vibration, airstream, force, temperature). The tactile display can be widely used because human skin includes a variety of receptors, namely mechanoreceptors (sensitive to pressure, vibration and slip), thermoreceptors (sensitive to temperature changes), nociceptors (responsible for pain) and proprioceptors (sensitive to position and movement), and any area of the human body can be activated by these receptors (e.g., finger or hand, wrist, arm, chest, back, forehead). Therefore, the tactile display has the potential for robust message delivery without generating additional cognitive demands or affecting driving performance. As shown in the tactile regulation of Table 8, current studies on the tactile regulation of driver's emotions in CAVs mainly focus on temperature and vibration. Temperature regulation can be used to create a comfortable and calming environment in the vehicle. For example, heated or cooled seats can provide physical comfort to drivers and help them manage stress levels. Vibration regulation can also be used to regulate drivers' emotional states. For example, vibration of the steering wheel or seat can provide physical feedback to help drivers stay alert and manage stress levels.

The study of temperature regulation by Schmidt et al. [154] reported that cool airstreams were associated with a decrease in low arousal and an increase in alertness and sympathetic activity, leading to better driving performance and acceptance. The other studies chose vibration to regulate drivers' emotions by instructing the driver to breathe and reported the effectiveness [150–153]. Besides, the study also found that haptic warning signals were as effective as audible warning signals in handling lane departures for both normal driving situations and driving secondary task situations [155]. In general, tactile regulation is a promising way for driver emotion regulation and can probably be more sufficiently utilized in the future.

### 6.1.3 Olfactory Regulation

Humans can also obtain information from olfaction. The olfaction has been proven to be good at activating the neural system, which can help drivers to perceive important information better. Hence, although less common, olfactory displays are slowly entering the field of automotive HMI as a new display for human-vehicle interaction [161]. Table 8 also investigates studies that have used different odors to regulate drivers' emotional states. Examples of odors studied include peppermint, rose, lavender, vanilla, cinnamon, and civet. Olfactory regulation can be used to create a calming and pleasant atmosphere in the vehicle.

For example, lavender has been found to have a calming effect on drivers.

Raudenbush et al. demonstrated that both peppermint and cinnamon decreased drivers' frustration and helped them focus on driving tasks [157]. Besides, Mustafa et al. [163] found that the presence of odors (vanilla and lavender) led drivers to experience positive emotions, such as relaxation with more feedback on comfortable and fresh feelings. Further, Dmitrenko et al. [156] selected the odors of different valence and arousal levels (e.g., rose, peppermint and civet) to regulate drivers' anger, and found that pleasant odors (rose and peppermint) could be able to shift drivers' emotions towards the positive valence. Furthermore, Dmitrenko et al. [159–161] proposed the method to map between different driving-related messages (“slow down,” “refuel,” and “pass by points of interest”) and four scents (lemon, lavender, mint, rose). Besides, some studies also found that intermittently applying scents can effectively keep drivers alert [158, 162]. Integrating olfactory regulation into CAVs can provide a more comprehensive and effective emotional regulation system for drivers. However, it is essential to note that olfactory responses can be subjective and vary between individuals. Therefore, a personalized approach to olfactory regulation may be necessary to ensure effectiveness for each individual driver. Overall, the olfactory regulation of driver's emotions in CAVs is a promising area of research, with the potential to significantly improve the driving experience for drivers in the future.

#### 6.1.4 Auditory Regulation

Besides the vision-based display, auditory displays are the second most common type of human-vehicle interaction. First, the auditory display is a high-resolution modality that can send complex information quickly through auditory icons (naturally occurring sounds), earcons (e.g., short tunes or bells), and spearcons (speech-based earcons). Second, the auditory display conveys directional (spatial) information, and auditory cues can be received from any direction. Third, compared to visual display, it reduces mental and visual distractions. Therefore, in the auditory regulation part of Table 8, studies on auditory emotion regulation are summarized, including adaptive music, empathetic speech, and auditory intervention. For example, calming music can be played in stressful driving situations to help drivers manage stress levels, and a virtual assistant can use empathetic speech to provide reassurance and guidance to drivers in unfamiliar driving situations.

Many studies envisage the application of adaptive music in the driving environment to regulate the driver's emotions [166, 167, 169]. Studies have indicated that recommending

calm music to drivers in a high arousal state could help them calm down, and will affect driver's emotional states and driving behavior in the long run [174, 178, 180]. However, other studies have pointed out that although music affected driving performance and reaction times, it had no significant effect on subjective emotion levels [168]. Empathetic speech requires empathetic adaptation of the in-vehicle voice agent's speech output to adapt to the emotional state of drivers and occupants in the driving environment. Studies on empathetic speech revealed that the matched speech arousal level could have a positive impact on driving performance and the frequency of interactions [175, 179]. Besides, the empathetic speech was liked by drivers, and could decrease the impact of sadness and anger on driving [170]. Empathetic speech may be a suitable method to regulate the driver's negative emotions. However, the above studies all worked with prototypical interactions. A clear description of empathy in voice interactions may require further research [16]. A typical form of auditory intervention is a reappraisal, which aims to illuminate frustrating situations in a more positive light. The study has shown that reappraisal may be an effective intervention that may nudge the driver towards a more positive state, thereby improving driving performance [172]. Another common form of auditory intervention is to instruct the driver to perform breathing exercises [152]. Other researchers have studied positive and negative comments [165], young and old voices [177], familiar and unfamiliar voices [171], warning voices [176], and different emotion regulation strategies to intervene driver emotions [181, 182]. In sum, auditory intervention may be an effective way to regulate the driver's emotions while rarely leading to distraction.

#### 6.2 V2X-Based Human Emotion Regulation

Seamless and real-time decisions of emotion regulation can be obtained in V2X, where drivers can be aware of their neighbors' emotions in the situated environment by communicating with intelligent traffic participants such as the cloud [183, 184] and neighboring vehicles [185].

In Ref. [183], a crowd-cloud collaborative and situation-aware music recommendation system is designed to diminish drivers' negative emotions. The sensing data and the driver's social information are collected and processed by the in-car smartphone. By collaborating with cloud computing, unified and seamless crowdsensing decisions can be made based on the fusion of information. This approach enables the in-vehicle smartphone to intelligently recommend preferable music to regulate drivers' negative emotions based on the situated environment by collaborating with its neighboring vehicles' phones and the cloud. Krishnan et al. [184] proposed a context-aware music emotion-mapping approach for drivers. The V2C communication is applied to aggregate the

Table 8 Summary of human emotion regulation in CAVs

M	Ref.	Emotion	No. Sub	Regulation Method	Platform	Measurement	Scenarios	Analysis method	Regulation Result
	[143]	Anger	21, male 12, female 9	Visual intervention, State feedback	Simulator	Self-reported scale, Physiological measurement, Driving behavior	Two-way four-lane highway scene, straight section	ANOVA	The regulation quality of the cold hue was better than the warm hue, the positive expression was better than the negative expression Drivers preferred to receive only safety-critical notifications of their state but appreciated a progressive status indicator for easier interpretation
	[144]	Anger	70, male 42, female 28	State feedback	Simulator	Self-reported scale	Three-lane highway with moderate traffic	Paired t-test	Drivers preferred to receive only safety-critical notifications of their state but appreciated a progressive status indicator for easier interpretation
Visual Regulation	[145]	Anger	30, male 15, female 15	Ambient light	Simulator	Self-reported scale, Physiological measurement	Multiple curves, intersections with stop lines and several sets of traffic lights	ANOVA	Blue light as a chromatic cue could promote relaxation and reduce physiological reactions associated with negative emotions
	[146]	Anger	52, male 19, female 33	Visual intervention	On-road	Self-reported scale, Driving behavior	The first route is no congestion on the trail, the second route includes major trunk roads and interstate sections	ANOVA	Advance warning of potential congestion reduced aggressive driving behavior of participants with high anger tendency, but increased aggressive behavior of participants with low anger tendency
	[144]	Anger	70, male 42, female 28	State feedback	Simulator	Self-reported scale	Three-lane highway with moderate traffic	Paired t-test	Drivers preferred to receive only safety-critical notifications of their state but appreciated a progressive status indicator for easier interpretation

Table 8 (continued)

M	Ref.	Emotion	No.Sub	Regulation Method	Platform	Measurement	Scenarios	Analysis method	Regulation Result
	[147]	Valence	328, male 219, female 109	Visual intervention, State feedback	Simulator	Self-reported scale	Near busy commercial street	ANOVA, paired t-test	Older respondents and less experienced drivers preferred systems with fewer visual elements. Extroverted drivers were more open to intervention
	[148]	Arousal	15, male 8, female 7	Visual relaxation techniques (VR)	On-road	Self-reported scale, Physiological measurement	Neighborhoods, mountains, and highway sections	ANOVA	Calm VR applications are ideal for automotive environments
	[149]	Valence, Arousal	12, male 8, female 4	Ambient light	Simulator	Self-reported scale, Physiological measurement, Driving behavior	Car-following task on a nearly empty highway, several annoying driving maneuvers from other drivers on a busy highway	ANOVA, Wilcoxon Signed Rank Test, Random Forest Classifier	Driving performance has been improved with the ambient lighting. Both blue and orange lights help drivers stay in lane
	[150]	Frustration	40, male 20, female 20	Vibration	On-road	Self-reported scale, Physiological measurement	A 12.3 mile GPS-guided route comprised of urban, mountain, and highway roads	Subject analysis	Most participants experienced the intervention as helpful to reduce breathing pace, and that this reduction led to a decrease in perceived stress
	[151]	Arousal	24, male 12, female 12	Vibration	On-road	Self-reported scale, Physiological measurement, Driving behavior	Simulate commute conditions, including turns, straight pathways, and stop signs	ANOVA	Haptic intervention to be successful in both user adoption and system effectiveness with an 82% rate of engagement in intervention and clear reduction of breathing rate and physiological arousal, with no effect on driving safety and minimal effect on performance

Table 8 (continued)

M	Ref.	Emotion	No.Sub	Regulation Method	Platform	Measurement	Scenarios	Analysis method	Regulation Result
	[152]	Arousal	24, male 12, female 12	Vibration	Simulator	Self-reported scale, Physiological measurement, Driving behavior	Cities (complex - users can interact with more cars, road signs, pedestrians, turns, etc.) and highways (simple)	Mixed analysis of variance paired t-test, post hoc multiple comparison	Tactile guidance systems can reduce the driver's respiratory rate and provide continuous post-intervention effect without affecting driving safety. Subjectively, most participants preferred tactile stimulation
	[153]	Arousal	10, male 8, female 2	Vibration	Simulator	Self-reported scale, Physiological measurement	Road (highway, city), times of day (day, night), and traffic levels (low, heavy, fast)	Subject analysis	Fast breathing interventions are more or less appropriate and beneficial (e.g., during long, monotonous drives on the highway or at night vs. complex or tense driving scenarios)
	[154]	Arousal	34, male 24, female 10	Temperature	Simulator	Self-reported scale, Physiological measurement, Driving behavior	Two monotonous highway route-seone with short-term cooling at the end and one without, with very little traffic	Two-sided Wilcoxon test	Cooling applied during a monotonous drive increased the alertness of driver, cooling has positive short-term effect on drivers' wakefulness
	[155]	Arousal	25	Vibration	Simulator	Driving behavior	A road test was conducted with a truck on a test track and in an urban environment	ANOVA	Haptic warning signals were as effective as the auditory warning signals in dealing with lane departures, both in normal driving situations and in a driving + secondary task situation

Table 8 (continued)

M	Ref.	Emotion	No.Sub	Regulation Method	Platform	Measurement	Scenarios	Analysis method	Regulation Result
	[156]	Anger	40, male 24, female 16	Rose, Peppermint, Civet	Simulator	Self-reported scale, Driving behavior	Straight country roads with a speed limit of 30 miles per hour, including green light, vehicle cut-off, slow front car, cut off the bicycle, etc	Wilcoxon signed ranks test, Kruskal Wallis test	Pleasant scents such as rose and mint are able to shift the driver's mood towards positive valence
	[157]	Frustration	25	Peppermint, Cinnamon	Simulator	Self-reported scale	One hour and two hours driving	Two-within analysis of variance	The use of cinnamon and mint will increase alertness, reduce time requirements and reduce frustration during driving scenes. The smell of mint reduces anxiety and fatigue
	[158]	Arousal	10, male 8, female 2	Mint, Rosemary, Eucalyptus, Lemon	Simulator	Driving behavior	Drive right along centerline on expressway with straight and curved sections at the velocity of 100 km/h	Lane departure analysis	Fragrance presentation to drivers at the wheel can induce wakefulness and peppermint presentation is suggested to be the most effective
	[159]	Arousal, Valence	22, male 14, female 8	Lavender, Mint, Lemon	Simulator	Self-reported scale, Driving behavior	Driving started on a motorway and participants were instructed to drive in any direction, following the traffic rules	Paired samples t-test	Olfactory notifications are perceived as less distracting, more comfortable, and more helpful than visual notifications. Drivers also make less driving mistakes when exposed to olfactory notifications
	[160]	Arousal, Valence	30, male 24, female 6	Lemon, Lavender, Mint, Rose	Simulator	Self-reported scale	Driving scenarios of "deceleration", "refueling" and "passing through the point of interest"	One way ANOVA	Establish a mapping between "lemon, peppermint" and "four different odors" (lemon, peppermint) and "one related to driving"

Table 8 (continued)

M	Ref.	Emotion	No. Sub	Regulation Method	Platform	Measurement	Scenarios	Analysis method	Regulation Result
	[161]	Arousal, Valence	21, male 11, female 10	Lavender	Simulator	Driving behavior	A speeding scenario, the driving task involves overtaking slower vehicles, there are also on coming vehicles and pedestrians involved	Two-way repeated measures ANOVA	When receiving olfactory notifications, participants have reduced the speed significantly faster, exceeded the speed limit fewer times
	[162]	Arousal, Valence	4, male 4	Perfume	Simulator	Self-reported scale, Driving behavior	Highway	Lateral displacement analysis	Intermittent presentation of the fragrance was effective in keeping a driver alert
	[163]	Arousal, Calm	10, male 5, female 5	Vanilla, Lavender	Simulator	Self-reported scale, Driving behavior	A road consisting of curves, straights, and roundabouts	Main effect analysis	The fragrance does not affect the reduction of speed. Due to the presence of aroma, the drivers' mood calmed down
	[5]	Anger	30, male 20, female 10	Empathetic speech	Empathetic speech	Self-reported scale, Physiological measurement, Driving behavior	Two-way four-lane highway scene, straight driving phase, obstacle avoidance phase	ANOVA	Active empathetic speech showed significantly better regulation quality than text-to-speech on some indicators of driver subjective feeling, prefrontal lobe activity and driving behavior
	[164]	Anger	60, male 34, female 26	Auditory intervention	Simulator	Self-reported scale, Driving behavior	The scenario included an urban road (with speed limit 40 mph) and a highway (with speed limit 50-65 mph)	ANOVA, paired sample t-test	Speech-based agents not only enhance driver situation awareness and driving performance, but also reduce their anger level and perceived workload

Table 8 (continued)

M	Ref.	Emotion	No.Sub	Regulation Method	Platform	Measurement	Scenarios	Analysis method	Regulation Result
	[165]	Anger	60, male 33, female 27	Auditory intervention	Simulator	Self-reported scale, Driving behavior	Free driving, lane change, pedestrians, traffic lights, crossroads, bus stop, a motorcycle came from the left crosswalk when turning right	ANOVA, post hoc analysis	Both positive and negative comments can reduce the driver's anger state and perceived workload, and improve driving performance, and positive comments are more effective than negative comments
	[166]	Anger	91, male 72, female 19	Adaptive music	Simulator	Self-reported scale, Physiological measurement, Driving behavior	Highways (easy cases) and rural roads have unexpected dangers, such as stones on the road (medium cases), roads in fog and snow (difficult cases)	ANOVA, paired t-test	Angry drivers who did not listen to music has riskier driving behavior than emotion-neutral drivers, music could help angry drivers react at the similar level to emotion-neutral drivers
	[167]	Anger	53, male 44, female 9	Adaptive music	Simulator	Self-reported scale, Driving behavior	A tunnel driving part with low visibility and a lane blockage, followed by an easily frustrating section full of frequent red lights and stop signs	One way ANOVA, paired t-test	Participants who listened to happy or sad music had significantly fewer driving errors than those who did not listen to music. No significant difference was found under happy and sad music conditions
	[168]	Anger	61, male 45, female 16	Adaptive music	Simulator	Self-reported scale, Physiological measurement, Driving behavior	Simple (roads), medium (rural roads) and difficult (fog and snow roads)	ANOVA, paired t-test, post hoc analysis	No specific emotional music significantly affected the participants' subjective emotional score. Angry drivers showed more aggressive driving behavior when listening to optional concerts

Table 8 (continued)

M	Ref.	Emotion	No.Sub	Regulation Method	Platform	Measurement	Scenarios	Analysis method	Regulation Result
	[169]	Anger	30, male 18, female 12	Adaptive music	Simulator	Self-reported scale, Physiological measurement, Driving behavior	Six lane highway with medium traffic flow. Participants were asked to follow the traffic in the central lane	ANOVA	Music with a specific rhythm or familiarity will be significantly better in alleviating anger and improving driving performance
Auditory Regulation	[170]	Anger, Sadness	60, male 42, female 18	Empathetic speech	Simulator	Self-reported scale, Physiological measurement, Driving behavior	Car following	ANOVA	Empathy assistants are the most positive way to resonate with users, and they work best to improve negative states
	[171]	Anger, Depression	60, male 30, female 30	Auditory intervention	Simulator	Self-reported scale, Driving behavior	Traffic congestion, traffic jams, pedestrian crossing and bad weather	Independent sample t-test paired sample t-test, two-way ANOVA	Familiar sounds has a more positive impact on the driver's driving feeling and driving performance (avoiding traffic accidents, abiding by traffic rules and lane deviation) than unfamiliar sounds
	[172]	Frustration	36, male 18, female 18	Auditory intervention (reappraisal)	Simulator	Self-reported scale, Driving behavior	Traffic congestion, vehicle confluence, traffic lights, Curved roads, slippery roads and low visibility	ANOVA	Participants in the reappraisal-down condition had better driving behavior and reported less negative emotions than participants in the other conditions
	[173]	Frustration	32, male 20, female 12	Situation selection Situation modification Attention deployment, Cognitive change, Response modulation	focus groups	Self-reported scale	-	Non-parametric tests	Speech interventions seem particularly promising, probably due to the high naturalness and the comparably little cognitive strain imposed by spoken language and auditory interaction on other tasks

Table 8 (continued)

M	Ref.	Emotion	No.Sub	Regulation Method	Platform	Measurement	Scenarios	Analysis method	Regulation Result
	[174]	Sadness, Happiness	17, male 8, female 9	Adaptive music	Simulator	Driving behavior	Highway, two lanes with each direction, use daytime dry road conditions with good visibility; the route does not include other road users	ANOVA	Happy music distracted drivers the most as their mean speed unexpectedly decreased and their lateral control deteriorated. Sad music influenced drivers in a different way as they drove slowly and kept their vehicle in its lane
	[175]	Happiness, Upset	40, male 20, female 20	Empathetic speech	Simulator	Self-reported scale, Driving behavior	Country roads (including other vehicles)	ANOVA	When user emotion matched car voice emotion (happy/energetic and upset/subdued), drivers had fewer accidents, attended more to the road (actual and perceived), and spoke more to the car
	[176]	Valence	36, male 18, female 18	Auditory intervention	Simulator	Self-reported scale, Driving behavior	Straight, curve, virtual passenger	ANOVA	Warnings associated with the environment works best. Drivers felt most at-ease, they liked the system, they rated the quality of the car higher, and their measured attention to the road was better
	[177]	Valence	18, male 9, female 9	Auditory intervention	Simulator	Self-reported scale, Driving behavior	Contains a variety of dangerous driving scenes: curves, thick fog, radar	ANOVA	There was a clear positive effect of driving with the in-car information system; drivers felt more confident driving, they completed the driving course in less time, and had fewer accidents

Table 8 (continued)

M	Ref.	Emotion	No.Sub	Regulation Method	Platform	Measurement	Scenarios	Analysis method	Regulation Result
	[178]	Valence	22, male 14, female 8	Adaptive music	On-road	Self-reported scale, Driving behavior	Three driving hours (morning, after- noon and evening) and three road types (resi- dential, Boule- vard and intercity Expressway)	ANOVA	In travel parameters (duration, distance, estimated travel speed) or perceived motion parameters (distraction, control and performance level), there was no difference between alternative music background and driver's preferred music
	[179]	Arousal	20	Empathetic speech	On-road	Driving behavior	Participants engaged in simulated cell phone calls while completing the primary task of driving while responding to the LEDs	ANOVA, paired t-test	Adding an angry emotional valence to the speech increased alertness level, resulting in reduced driver distraction, likely via increases in right frontopari- etal networks
	[152]	Arousal	24, male 12, female 12 (breath)	Auditory interven- tion (breath)	Simulator	Self-reported scale, Physiological meas- urement, Driving behavior	Cities (complex - users can interact with more cars, road signs, pedestrians, turns, etc.) and highways (simple)	Mixed analysis of variance paired t-test, post hoc multiple comparison	Voice guidance sys- tems can reduce the driver's respiratory rate and provide continuous post- intervention effect without affecting driving safety
	[180]	Valence, Arousal	19, male 13, female 6	Adaptive music	Simulator	Self-reported scale, Physiological meas- urement, Driving behavior	Rural roads, simple curves (about 80%), with an average of 15 buses passing through every minute	ANOVA, post hoc multiple comparison	Music affects the mood when driving, respiration rate was lower during music listening compared to rides without music, and the effect of music on heart rate is not found

crowd-sensed music emotion-mapping data to improve the efficiency and effectiveness of the music delivery for drivers. By incorporating the social context information of a driver (e.g., gender, age, and personality) collected from the cloud, the vehicle's system ensures the timely delivery of suitable music according to different driver emotions. A location-based method was designed in Ref. [185] that enables road users to send and receive feedback on their driving behaviors through V2V communication. On the one hand, a driver can express his appreciation and disapproval towards nearby drivers about their polite and impolite driving behavior by gestures. On the other hand, he can receive others' evaluations through audio and visual feedback. With the help of a driving simulator, the system is demonstrated to exert a positive influence on driving behaviors.

### 6.3 Regulation Scenarios

Experimental scenarios are particularly important for design, verification, and validation efforts. Scenes in the real world are infinitely rich, extremely complex, and unpredictable. It is very difficult to reproduce these scenes entirely in a virtual environment. Therefore, the key to realizing the test and verification of the driving simulator is to use the limited test scene to map the infinitely rich world. The elements of the test scene include two parts: the elements of test vehicles and the elements of the external traffic environment. The elements of the external traffic environment also include static environment elements, dynamic environment elements, traffic participants elements, and meteorological elements.

Most previous studies employed simulated scenarios to regulate human emotions. In Table 8, highway was a common experiment scenario and some studies simulated the highway scene to study visual intervention [143, 144, 148], auditory intervention [166, 169, 174, 178, 181], vibration intervention [152–155], and olfactory intervention [158, 159, 161, 162]. In addition, to make the scenario closer to real driving, some studies designed the car following [149, 170], curves, intersections [145], urban roads [152], and traffic signs, traffic lights, traffic density [166], traffic participants [171, 177] factors. Some researchers also conducted simulated driving experiments in different time periods (i.e., morning, afternoon, and evening) [153, 178]. A small number of on-road studies have been carried out for emotion regulation. Some researchers chose country trails and highways to study the effect of the visual intervention on the regulation [146]. Others studied the effect of haptic regulation using commuter routes that contain straights, curves, and traffic signs [151]. In sum, researchers can choose highways, complex urban roads, and simulated driving environments such as weather and driving tasks according to specific requirements when designing scenarios.

### 6.4 Regulation Quality Analysis Method

The analysis methods for regulation quality mainly include measurements and analysis. As shown in Table 8, driver emotion regulation in CAVs is commonly measured using self-report scales, driving behavior, and physiological measurements such as heart rate and respiration. Self-report scales are the most commonly used method to measure regulation quality, while some studies also investigate driving behavior and physiological changes under different regulation methods. The purpose of emotion regulation analysis is to examine differences in regulation quality across different regimens. Due to the different data distributions, existing studies have adopted different data analysis methods. Most studies have used ANOVA for quality analysis of emotion regulation due to its efficiency in differences analysis. Post hoc analysis is often performed after ANOVA to analyze differences between different regulation schemes. Other studies have used paired t-tests, independent-samples t-tests, main effects analysis, and non-parametric tests for regulation quality analysis.

- *Summary of human emotion regulation in CAVs:* This section discusses methods for in-cabin human emotion regulation, V2X-based human emotion regulation, regulation scenarios, and regulation quality analysis method. In terms of emotion regulation schemes, the existing research is reviewed from the aspects of visual regulation, auditory regulation, tactile regulation, and olfactory regulation. Existing studies mainly focus on the driver's single-modal emotion regulation, such as ambient light, adaptive music, vibration and odor. Benefiting the development of intelligent cockpit interaction technology in CAVs, studies on multimodal emotion regulation schemes for drivers and occupants are also potential.

## 7 Discussion and Perspectives

Human emotion is the main research direction in CAVs development. To improve acceptance, safety, and comfort in CAVs, an affective connected automated driving (ACAD) is proposed based on the previously reviewed literature in this paper. Figure 5 shows ACAD framework for human emotion research based on the human-vehicle-road coupling process in the context of CAVs. As shown, the proposed ACAD framework consists of three main layers, namely the human layer, the automated vehicle layer, and the road layer. The human layer describes the human's emotional generation process and human behaviors in CAV. When the human is in a specific situation and is stimulated by the context, the human will have a cognitive appraisal for the current situation. This cognitive appraisal will then lead to different

emotional responses, including subjective experiences, behavioral responses, and physiological changes that result in different emotional states and human behavior in CAVs.

The automated vehicle layer presents three modules of affectively automated driving, which are in-cabin human emotion detection, driving context perception, and emotion-based automated driving. In the in-cabin human emotion detection, both the physiological and physical behavioral data of the occupants is collected with multimodal sensor systems in the cabin. The data-driven multimodal emotion detection module output the emotion detection results based on the multimodal sensing data processing and information fusion. Meanwhile, in the driving context perception module, the driving scenario information, including obstacle positions, lane marks and vehicle states are perceived. While serving as the important input of automated driving, the traffic context is the key factor in inducing human emotions for better inference. Last, in the emotion-based automated driving module, the driving risk assessment will consider the driving scenario information and the in-cabin human emotion comprehensively. With the help of emotion detection, risk assessment can be more advanced and precise. Based on the risk, the decision-making can be generated by integrating the traffic rules and emotion regulation rules, which will further control vehicles according to different vehicle actuation configurations and fed back to the driver according to multimodal interaction interfaces, respectively.

The road layer includes vehicle communication with the surroundings, online driver emotion detection, history data collection, and the creation of emotion regulation strategies. The vehicle communicates with the surrounding mainly consisting of V2V, V2I, V2C and vehicle to pedestrian (V2P). Online driver emotion detection includes data processing, data fusion, and the output of driver emotion. The basic biometrics, habitual driving behavior and other drivers' information are stored in the history data. Besides, online driver emotion detection and history data can be employed to generate driver emotion regulation strategies for different drivers to guide emotion-based automated driving. In sum, the proposed ACAD framework will dynamically record multimodal data for real-time emotion detection and further produce adaptive emotion regulation strategies to improve the affective interaction quality between humans and CAVs and promote driving safety.

Possible future research directions of human emotion in CAVs may consist of several aspects as follows:

### 7.1 Human Emotion Dataset in Driving

Over the past decade, the driver's emotion datasets during driving proposed by some researchers have greatly promoted the research of human emotion detection in driving [9, 87,

186]. In the future, researchers still need rich and repeatable experimental datasets. To collect a reliable human emotion dataset in driving, the following issues should be considered.

- *Driver emotion dataset: Induction:* Since driving is a complicated and long-lasting task, the intensity duration of the induced emotional state and the naturalness degree of induction are essential to the dataset and have a more significant impact on driving performance. *Measurements:* The dataset should include high-quality measurements of essential clues related to human emotional experience and expression (face, speech, physiology, behavior) while considering the low intrusiveness of the measurement metrics, which may include some indirect and non-intrusive indicators. *Annotation and scenarios:* The dataset should include annotations of the driver's emotional states, and the dataset should consider the driving scenario and duration, and collect data as much as possible while driving the actual vehicle in real traffic. *Characteristics:* The personality characteristics of the participants, including age, gender, driving experience, and driving style, should also be considered.
- *Human emotion interaction dataset: In-cabin interaction between occupants:* Future datasets should consider in-cabin interaction between occupants, as the emotional component plays a major role in human-object interactions in vehicles and interactions between occupants. *Driver interaction with other road users:* Studying the interaction behavior of humans with other road users under emotion will help to study different driving strategies for CAVs and improve driving safety. *Interactions between emotions and different states:* Studying the interaction between human emotions and motion sickness, load, and distraction in driving will help to improve the emotional experience of intelligent mobile spaces. Interaction datasets should also consider factors of induction, measurement, annotation, and characteristics.

### 7.2 Human Emotion Detection in CAVs

Human emotion detection in CAVs is a research field that is increasingly growing in importance and attention. To stimulate research in this area, previous research is analyzed to discuss some future research opportunities.

- *Multimodal driver emotion detection:* Various measurement methods and algorithms (physiology, behavior, subjective scales) have been used in the field of driver emotion detection. However, unimodal driver emotion detection does not always accurately recognize the internal state (e.g., facial expressions do not always map exactly or via a simply fixed mapping onto the internal state). A more comprehensive multimodal detection

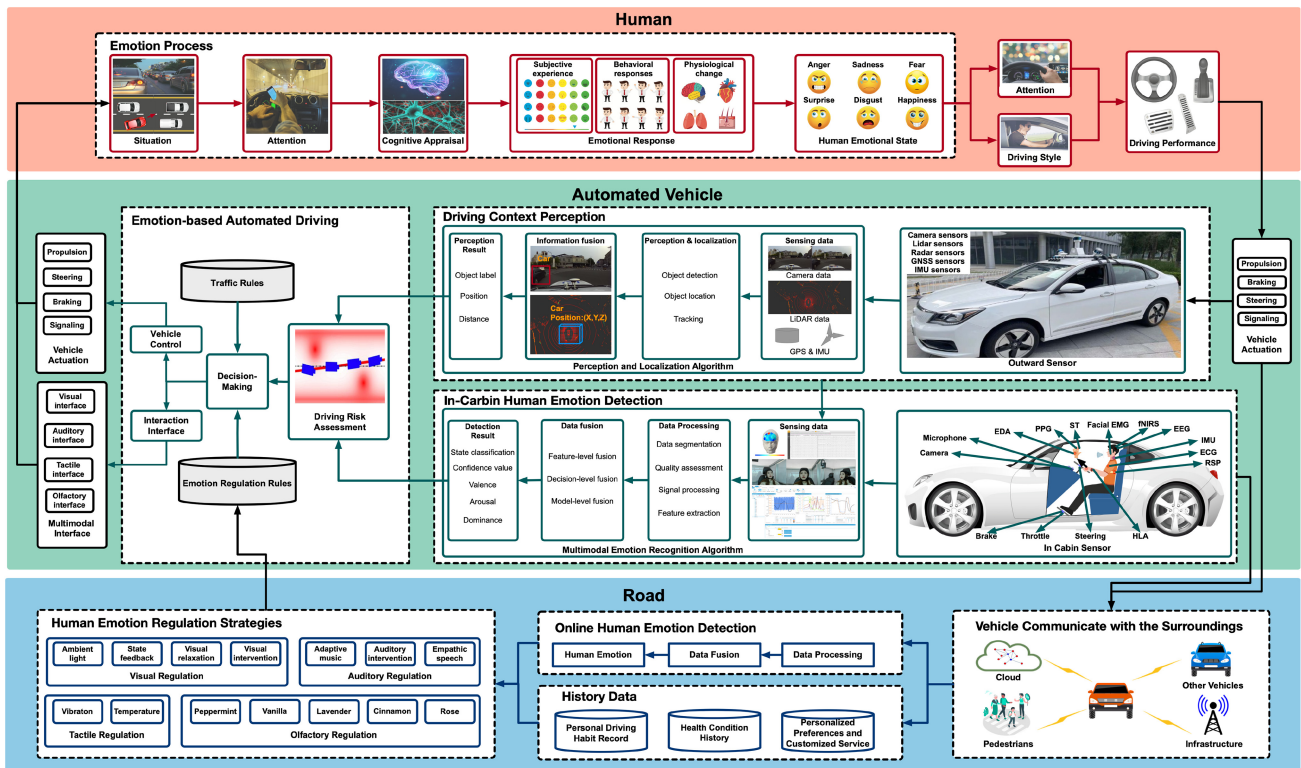


Fig. 5 Framework of affective connected automated driving (ACAD)

method that combines multiple signals can better evaluate driver emotions. Notably, real-time communication of multiple sensors and fusion of metrics in an emotional multimodal recognition environment is critical.

- *Facial anonymization in emotion detection:* Facial expression is commonly used in emotion detection among current studies. However, this detection method may cause a privacy leak in the future if implementing such an emotion detection algorithm to the cabin. Therefore, it will be important to balance facial anonymization and important information when developing the algorithm. Using a model, such as GAN network, to create a virtual face might be a solution.
- *Occupants' emotion detection:* The studies on occupants' emotion detection is in progress. Some researchers applied the algorithms that were developed for general emotion detection to the occupants' emotion detection. Although the emotions of occupants might be similar to the general emotions since there are no driving tasks involved. However, further studies are needed to prove the similarity between general emotion detection and occupants' emotion detection in the driving context. Besides, considering the continuous characteristic of human emotions, occupant emotion detection needs to be discussed from dimensional emotion (valence, arousal, dominance) in the future for CAVs.

- *Online driver emotion detection:* At present, most research for automatic emotion recognition in vehicles is conducted by processing the data offline because different methods can be compared to a fixed set of data. However, as more and more sophisticated sensors are integrated into the vehicle, and a more massive amount of data is collected in real-time, driver emotion detection requires excellent storage space and computing functions to support the data processing and fusion. Therefore, V2X-based online driver emotion detection systems will play an essential role in the future, which can effectively leverage information when it is most needed.
- *Driving-context-based detection:* Previous studies have used experimental scenarios to trigger specific emotions and assume that certain scenarios induce specific emotional states, so it is crucial to identify the source or events of emotional changes (e.g., turning, overtaking, and road congestion). As current vehicles are increasingly equipped with sensors that provide rich contextual information, a more comprehensive driver emotion detection method will improve their recognition performance by considering the integration of the vehicle, the driver and the environment.
- *Combine with cognitive appraisal:* Cognitive appraisal is essential for emotion generation. Neuroscience research has shown that the generation of emotions is closely

related to the activation of different regions in the brain. Therefore, based on neuroscience research, the study of the cognitive appraisal process may help emotion researchers better detect driver emotions.

- *Driver emotion detection in different cultures:* Since different cultures have different expressions of emotions (e.g., the expression of facial expressions differs between East Asian and Western people), this circumstance needs to be considered in driver emotion detection. According to the different cultural backgrounds of drivers, different methods should be explored to improve the accuracy of emotion recognition.

### 7.3 Human Emotion Regulation in CAVs

The application of driver emotion regulation technologies in the automotive industry field may appear on a large scale within the next 5-10 years. Combined with the summary in Section 6, several potential aspects of human emotion regulation are discussed.

- *Parameters emotion regulation:* Since most of the current studies work with prototypical interactions, making it not easy to describe a specific displayed emotion regulation effect clearly. Therefore, it is necessary to investigate the emotion regulation effect of different parameters in different displays (e.g., visual regulation: color, shape, expression; tactile regulation: activation time, frequency, intensity, or amplitude). Such outcomes will provide automotive HMI practitioners with knowledge on how to use different displays to design emotions in the vehicle. Besides, the relationship between these parameters and driving distraction should also be considered.
- *Multimodal emotion regulation:* The driving context is crucial to the study of emotion regulation. Unimodal emotion regulation may be ineffective due to rapid changes in driving scenarios (e.g., strong sunlight reflections can invalidate visual regulation; bumpy roads can mask tactile regulation; noisy construction sites might make auditory regulation be drowned out). Therefore, the driver's emotion regulation should be multimodal and adapt to the context.
- *Emotion regulation for different cultures and characteristics:* Because emotions are culturally sensitive, this factor needs to be considered in the automotive HMI. Also, drivers' personality characteristics (e.g., extraversion or introversion) and emotional regulation strategies (reappraisal and suppression) preferences should also be considered. According to different cultures and characteristics of drivers, the corresponding emotional regulation methods should be explored.
- *Driver emotion regulation from the V2X:* Because driver characteristics play an important role in emotion regula-

tion, driver emotion regulation needs to consider a large amount of historical characteristics data (e.g., human driving habits, health history, and personalized preferences). By integrating the driver's social background and historical information collected from the V2X, the vehicle's emotion regulation system can ensure that appropriate regulation strategies are provided in time according to different driving contexts.

- *Privacy:* With the rapid development of CAVs, in future scenarios, the vehicle will become a relaxed and social interaction environment (e.g., a living room or a workplace). How to respect the privacy of the occupants in the vehicle while detecting and regulating emotions will be a vital research topic.

## 8 Conclusions

Affective human-vehicle interaction in the intelligent cockpit is a research field that is increasingly growing in importance and attention with the development of CAVs. To help motivate research in this field, this paper first emphasizes the potential applications and related research areas of affective human-vehicle interactions in CAVs, as well as introduces the definition and classification of emotions. This paper then summarizes the relationship between different emotions and emotional expressions from four aspects: facial expression, speech, body gesture, and physiological changes. The paper also introduces the human emotion experiment in driving by introducing the detailed methodologies in experiment platform, induction, measurement, and annotation. Moreover, this paper discusses human emotion detection in CAVs by introducing the detailed methodologies in in-cabin driver emotion detection, in-cabin occupant emotion detection, driving-context-based driver emotion detection and V2X-based human emotion detection. This paper also studies human emotion regulation in CAVs consisting of in-cabin regulation methods, V2X-based human emotion regulation, regulation scenarios, and regulation quality analysis methods. Finally, this paper proposed the ACAD framework and discusses future trends from three aspects: human emotion dataset in driving, human emotion detection in CAVs and human emotion regulation in CAVs to improve acceptance, safety, comfort and enjoyment levels for users CAVs.

**Acknowledgements** This work was supported by Natural Science Foundation of China (52302497, 52272420)

**Author Contributions** W. L.: Conceptualization, Formal analysis, Writing—original draft, Writing—review & editing. G. L.: Writing—original draft. R. T.: Writing - review & editing. C.W.: Writing—review & editing. Z. S.: Writing—review & editing. Y. L.: Conceptualization. G. G.: Supervision, Resources. D. C.: Supervision, Resources. K. L.: Supervision, Resources.

## Declarations

**Conflict of interest** On behalf of all authors, the corresponding author states that there is no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Peng, L., Wang, H., Li, J.: Uncertainty evaluation of object detection algorithms for autonomous vehicles. *Autom. Innov.* **4**(3), 241–252 (2021)
- MIITC: Roadmap of intelligent and connected vehicle 2.0. [http://www.gov.cn/xinwen/2020-11/19/content\\_5562464.htm](http://www.gov.cn/xinwen/2020-11/19/content_5562464.htm) (2020)
- Cui, M., Huang, H., Xu, Q., et al.: Survey of intelligent and connected vehicle technologies: Architectures, functions and applications. *J. Tsinghua Univ. Sci. Tech.* **62**(3), 493–508 (2022)
- Peng, B., Sun, Q., Li, S.E., Kum, D., Yin, Y., Wei, J., Gu, T.: End-to-end autonomous driving through dueling double deep q-network. *Autom. Innov.* **4**(3), 328–337 (2021)
- Li, W., Wu, L., Wang, C., Xue, J., Hu, W., Li, S., Guo, G., Cao, D.: Intelligent cockpit for intelligent vehicle in metaverse: a case study of empathetic auditory regulation of human emotion. *IEEE Trans. Syst. Man Cybern. Syst.* **53**(4), 2173–2187 (2022)
- ISO/TR-21959-1: Road vehicles - Human performance and state in the context of automated driving - Part 1: Common underlying concepts. <https://www.iso.org/standard/78088.html> (2020)
- Nunes, A., Reimer, B., Coughlin, J.F.: People must retain control of autonomous vehicles. Nature Publishing Group (2018)
- Quante, L., Zhang, M., Preuk, K., Schiebl, C.: Human performance in critical scenarios as a benchmark for highly automated vehicles. *Autom. Innov.* **4**(3), 274–283 (2021)
- Li, W., Tan, R., Xing, Y., Li, G., Li, S., Zeng, G., Wang, P., Zhang, B., Su, X., Pi, D., et al.: A multimodal psychological, physiological and behavioural dataset for human emotions in driving tasks. *Sci. Data* **9**(1), 1–20 (2022)
- Terken, J., Pflöging, B.: Toward shared control between automated vehicles and users. *Autom. Innov.* **3**(1), 53–61 (2020)
- Clark, J.R., Stanton, N.A., Revell, K.: Automated vehicle handover interface design: Focus groups with learner, intermediate and advanced drivers. *Autom. Innov.* **3**(1), 14–29 (2020)
- Li, S., Zhang, T., Liu, N., Zhang, W., Tao, D., Wang, Z.: Drivers' attitudes, preference, and acceptance of in-vehicle anger intervention systems and their relationships to demographic and personality characteristics. *Int. J. Ind. Ergon.* **75**, 102899 (2020)
- Chen, L., Wu, M., Zhou, M., Liu, Z., She, J., Hirota, K.: Dynamic emotion understanding in human-robot interaction based on two-layer fuzzy svr-ts model. *IEEE Trans. Syst. Man Cybern. Syst.* **50**(2), 490–501 (2020)
- Liang, Y., Li, Y., Yu, Y., Zhang, Z., Zheng, L., Ren, Y.: Path-following control of autonomous vehicles considering coupling effects and multi-source system uncertainties. *Autom. Innov.* **4**(3), 284–300 (2021)
- Zepf, S., Hernandez, J., Schmitt, A., Minker, W., Picard, R.W.: Driver emotion recognition for intelligent vehicles: a survey. *ACM Comput. Surv.* **53**(3), 1–30 (2020)
- Braun, M., Weber, F., Alt, F.: Affective automotive user interfaces—reviewing the state of emotion regulation in the car. *arXiv: 2003.13731* (2020)
- Gross, J.J.: *Handbook of Emotion Regulation*. Guilford Press, New York (2013)
- Ekman, P.: An argument for basic emotions. *Cogn. Emot.* **6**(3–4), 169–200 (1992)
- Plutchik, R.: The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *Am. Sci.* **89**(4), 344–350 (2001)
- Posner, J., Russell, J.A., Peterson, B.S.: The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Dev. Psychopathol.* **17**(3), 715–734 (2005)
- Mehrabian, A.: Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament. *Curr. Psychol.* **14**(4), 261–292 (1996)
- Lazarus, R.S., Lazarus, R.S.: *Emotion and Adaptation*. Oxford University Press, Oxford (1991)
- Ekman, P., Friesen, W.V., O'sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W.A., Pitcairn, T., Ricci-Bitti, P.E.: Universals and cultural differences in the judgments of facial expressions of emotion. *J. Pers. Soc. Psychol.* **53**(4), 712 (1987)
- Lazarus, R.S.: Progress on a cognitive-motivational-relational theory of emotion. *Am. Psychol.* **46**(8), 819 (1991)
- Kleinginna, P.R., Kleinginna, A.M.: A categorized list of emotion definitions, with suggestions for a consensual definition. *Motiv. Emot.* **5**(4), 345–379 (1981)
- Ekman, P., Friesen, W.V.: Constants across cultures in the face and emotion. *J. Pers. Soc. Psychol.* **17**(2), 124 (1971)
- Russell, J.A.: A circumplex model of affect. *J. Pers. Soc. Psychol.* **39**(6), 1161 (1980)
- Russell, J.A.: Core affect and the psychological construction of emotion. *Psychol. Rev.* **110**(1), 145 (2003)
- Barrett, L.F., Mesquita, B., Ochsner, K.N., Gross, J.J.: The experience of emotion. *Annu. Rev. Psychol.* **58**, 373–403 (2007)
- Zeng, Z., Pantic, M., Roisman, G.I., Huang, T.S.: A survey of affect recognition methods: audio, visual, and spontaneous expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(1), 39–58 (2008)
- Keltner, D., Sauter, D., Tracy, J., Cowen, A.: Emotional expression: advances in basic emotion theory. *J. Nonverbal Behav.* **43**(2), 133–160 (2019)
- Ekman, P., Friesen, W.V., Hager, J.C.: *Facial Action Coding System: the Manual*. Research Nexus, Salt Lake City, Utah (2002)
- Du, S., Tao, Y., Martinez, A.M.: Compound facial expressions of emotion. *Proc. Natl. Acad. Sci. U.S.A.* **111**(15), 1454–1462 (2014)
- Russell, J.A., Bachorowski, J.-A., Fernández-Dols, J.-M.: Facial and vocal expressions of emotion. *Annu. Rev. Psychol.* **54**(1), 329–349 (2003)
- Kanade, T., Cohn, J.F., Tian, Y.: Comprehensive database for facial expression analysis. In: *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 46–53 (2000). IEEE
- Russell, J.A.: Is there universal recognition of emotion from facial expression: a review of the cross-cultural studies. *Psychol. Bull.* **115**(1), 102–141 (1994)
- Hess, U., Banse, R., Kappas, A.: The intensity of facial expression is determined by underlying affective state and social situation. *J. Pers. Soc. Psychol.* **69**(2), 280–288 (1995)

38. Kring, A.M., Sloan, D.M.: The facial expression coding system (FACES): development, validation, and utility. *Psychol. Assess.* **19**(2), 210 (2007)
39. Zeng, Z., Pantic, M., Roisman, G.I., Huang, T.S.: A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(1), 39–58 (2009)
40. Whissell, C.M.: The dictionary of affect in language. In: *The Measurement of Emotions*, pp. 113–131. Elsevier, Amsterdam (1989)
41. Neumann, M., et al.: Cross-lingual and multilingual speech emotion recognition on english and french. In: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5769–5773 (2018). IEEE
42. Latif, S., Qayyum, A., Usman, M., Qadir, J.: Cross lingual speech emotion recognition: Urdu vs. western languages. In: *2018 International Conference on Frontiers of Information Technology (FIT)*, pp. 88–93 (2018). IEEE
43. Bachorowski, J.-A., Owren, M.J.: Vocal expressions of emotion. *Handbook of emotions* **3**, 196–210 (2008)
44. Banse, R., Scherer, K.R.: Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* **70**(3), 614 (1996)
45. Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., Taylor, J.G.: Emotion recognition in human-computer interaction. *IEEE Signal Process. Mag.* **18**(1), 32–80 (2001)
46. Juslin, P.N., Laukka, P.: Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion* **1**(4), 381–412 (2001)
47. Bachorowski, J.-A.: Vocal expression and perception of emotion. *Curr. Dir. Psychol. Sci.* **8**(2), 53–57 (1999)
48. Bachorowski, J.-A., Owren, M.J.: Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context. *Psychol. Sci.* **6**(4), 219–224 (1995)
49. Karadoğan, S.G., Larsen, J.: Combining semantic and acoustic features for valence and arousal recognition in speech. In: *Int. Workshop Cog. Inf. Process.*, pp. 1–6 (2012). IEEE
50. Noroozi, F., Corneanu, C.A., Kamińska, D., Sapiński, T., Escalera, S., Anbarjafari, G.: Survey on emotional body gesture recognition. *IEEE Trans. Affect. Comput.* **12**(2), 505–523 (2018)
51. Ruthrof, H.: *The Body in Language*. Bloomsbury Publishing, Bloomsbury (2015)
52. Molchanov, P., Gupta, S., Kim, K., Kautz, J.: Hand gesture recognition with 3d convolutional neural networks. In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 1–7 (2015)
53. Pease, B., Pease, A.: *The Definitive Book of Body Language: The Hidden Meaning Behind People's Gestures and Expressions*. Bantam, New York (2008)
54. Dael, N., Mortillaro, M., Scherer, K.R.: The body action and posture coding system (bap): development and reliability. *J. Nonverbal Behav.* **36**(2), 97–121 (2012)
55. Coulson, M.: Attributing emotion to static body postures: recognition accuracy, confusions, and viewpoint dependence. *J. Nonverbal Behav.* **28**(2), 117–139 (2004)
56. Gunes, H., Piccardi, M.: Affect recognition from face and body: early fusion vs. late fusion. In: *2005 IEEE International Conference on Systems, Man and Cybernetics*, vol. 4, pp. 3437–3443 (2005). IEEE
57. Castellano, G., Villalba, S.D., Camurri, A.: Recognising human emotions from body movement and gesture dynamics. In: *Int. Conf. Affect. Comput. Intell. Interact.*, pp. 71–82 (2007). Springer
58. Saha, S., Datta, S., Konar, A., Janarthanan, R.: A study on emotion recognition from body gestures using kinect sensor. In: *2014 international conference on communication and signal processing*, pp. 056–060 (2014). IEEE
59. Glowinski, D., Mortillaro, M., Scherer, K., Dael, N., Volpe, G., Camurri, A.: Towards a minimal representation of affective gestures. In: *2015 international conference on affective computing and intelligent interaction (ACII)*, pp. 498–504 (2015). IEEE
60. Kosti, R., Alvarez, J.M., Recasens, A., Lapedriza, A.: Emotion recognition in context. In: *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 1667–1675 (2017)
61. Rani, P., Liu, C., Sarkar, N., Vanman, E.: An empirical study of machine learning techniques for affect recognition in human-robot interaction. *Pattern Anal. Appl.* **9**(1), 58–69 (2006)
62. Öhman, A., Hamm, A., Hugdahl, K.: Cognition and the autonomic nervous system: orienting, anticipation, and conditioning. *Handbook of psychophysiology*, 533–575 (2000)
63. Williams, C.E., Stevens, K.N.: Vocal correlates of emotional states. *Speech evaluation in psychiatry*, 221–240 (1981)
64. Van Der Vloed, G., Berentsen, J.: Measuring emotional wellbeing with a non-intrusive bed sensor. In: *IFIP Conf. Hum. Comput. Interact.*, pp. 908–911 (2009). Springer
65. Norman, G.J., Berntson, G.G., Cacioppo, J.T.: Emotion, somatovisceral afference, and autonomic regulation. *Emot. Rev.* **6**(2), 113–123 (2014)
66. Stephens, C.L., Christie, I.C., Friedman, B.H.: Autonomic specificity of basic emotions: evidence from pattern classification and cluster analysis. *Biol. Psychol.* **84**(3), 463–473 (2010)
67. Kreibitz, S.D.: Autonomic nervous system activity in emotion: a review. *Biol. Psychol.* **84**(3), 394–421 (2010)
68. Calvo, R.A., D'Mello, S.: Affect detection: an interdisciplinary review of models, methods, and their applications. *IEEE Trans. Affect. Comput.* **1**(1), 18–37 (2010)
69. Lang, P.J., Greenwald, M.K., Bradley, M.M., Hamm, A.O.: Looking at pictures: affective, facial, visceral, and behavioral reactions. *Psychophysiology* **30**(3), 261–273 (1993)
70. Cacioppo, J.T., Berntson, G.G., Larsen, J.T., Poehlmann, K.M., Ito, T.A.: The psychophysiology of emotion. *Handbook of emotions* **2**, 173–191 (2000)
71. Panksepp, J.: Neurologizing the psychology of affects: how appraisal-based constructivism and basic emotion theory can coexist. *Perspect. Psychol. Sci.* **2**(3), 281–296 (2007)
72. Murphy, F.C., Nimmo-Smith, I.A.N., Lawrence, A.D.: Functional neuroanatomy of emotions: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* **3**(3), 207–233 (2003)
73. Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., Barrett, L.F.: The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* **35**(3), 121–143 (2012)
74. Pace-Schott, E.F., Amole, M.C., Aue, T., Balconi, M., Bylsma, L.M., Critchley, H., Demaree, H.A., Friedman, B.H., Gooding, A.E.K., Gosseries, O., et al.: Physiological feelings. *Neurosci. Biobehav. Rev.* **103**, 267–304 (2019)
75. Damasio, A.R., Grabowski, T.J., Bechara, A., Damasio, H., Ponto, L.L., Parvizi, J., Hichwa, R.D.: Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nat. Neurosci.* **3**(10), 1049–1056 (2000)
76. Vytal, K., Hamann, S.: Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis. *J. Cognit. Neurosci.* **22**(12), 2864–2885 (2010)
77. Kirby, L.A.J., Robinson, J.L.: Affective mapping: an activation likelihood estimation (ALE) meta-analysis. *Brain Cogn.* **118**, 137–148 (2017)
78. Bella, F.: Driver perception hypothesis: Driving simulator study. *Transp. Res. Pt. F-Traffic Psychol. Behav.* **24**, 183–196 (2014)
79. De Winter, J., Van Leeuwen, P., Happee, R.: Advantages and disadvantages of driving simulators: A discussion. In: *Proc. Meas. Behav.* (2012)

80. Yu, J., Tseng, P., Muggleton, N.G., Juan, C.-H.: Being watched by others eliminates the effect of emotional arousal on inhibitory control. *Front. Psychol.* **6**, 4 (2015)
81. Schweitzer, J., Green, P.: Task acceptability and workload of driving city streets, rural roads, and expressways: Ratings from video clips. Univ. of Michigan, Ann Arbor, *Transp. Res. Inst. Technical report* (2007)
82. Sukhvasi, S.B., Sukhvasi, S.B., Elleithy, K., El-Sayed, A., Elleithy, A.: A hybrid model for driver emotion detection using feature fusion approach. *Int. J. Environ. Res. Public Health* **19**(5), 3085 (2022)
83. Azizan, M.A., Fatih, M.I.A., Nabila, A., Norhashim, N., Omar, M.N.: Development of real-time emotion recognition system based on machine learning algorithm. In: *Human-Centered Technol. for a Better Tomorrow*, pp. 101–114. Springer, New York (2022)
84. Xiao, H., Li, W., Zeng, G., Wu, Y., Xue, J., Zhang, J., Li, C., Guo, G.: On-road driver emotion recognition using facial expression. *Appl. Sci.* **12**(2), 807 (2022)
85. Xia, Q., Li, J., Dong, A.: Road rage recognition system based on face detection emotion. In: *Int. Conf. Broadband Commun. Networks Syst.*, pp. 174–181 (2021). Springer
86. Li, W., Zeng, G., Zhang, J., Xu, Y., Xing, Y., Zhou, R., Guo, G., Shen, Y., Cao, D., Wang, F.-Y.: Cogemonet: A cognitive-feature-augmented driver emotion recognition model for smart cockpit. *IEEE Trans. Comput. Social Syst.* **9**(3), 667–678 (2022)
87. Li, W., Cui, Y., Ma, Y., Chen, X., Li, G., Zeng, G., Guo, G., Cao, D.: A spontaneous driver emotion facial expression (defe) dataset for intelligent vehicles: Emotions triggered by video-audio clips in driving scenarios. *IEEE Trans. Affect. Comput.* **14**(1), 747–760 (2020)
88. Cui, Y., Ma, Y., Li, W., Bian, N., Li, G., Cao, D.: Multi-emonet: a novel multi-task neural network for driver emotion recognition. *IFAC-PapersOnLine* **53**(5), 650–655 (2020)
89. Mălăescu, A., Duțu, L.C., Sultana, A., Filip, D., Ciuc, M.: Improving in-car emotion classification by nir database augmentation. In: *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*, pp. 1–5 (2019). IEEE
90. Lee, K.W., Yoon, H.S., Song, J.M., Park, K.R.: Convolutional neural network-based classification of driver's emotion during aggressive and smooth driving using multi-modal camera sensors. *Sensors* **18**(4), 957 (2018)
91. Ma, Z., Mahmoud, M., Robinson, P., Dias, E., Skrypchuk, L.: Automatic detection of a driver's complex mental states. In: *Computational Science and Its Applications—ICCSA 2017: 17th International Conference, Trieste, Italy, July 3–6, 2017, Proceedings, Part III 17*, pp. 678–691 (2017). Springer
92. Gao, H., Yüce, A., Thiran, J.-P.: Detecting emotional stress from facial expressions for driving safety. In: *2014 IEEE international conference on image processing (ICIP)*, pp. 5961–5965 (2014). IEEE
93. Requardt, A.F., Ihme, K., Wilbrink, M., Wendemuth, A.: Towards affect-aware vehicles for increasing safety and comfort: recognising driver emotions from audio recordings in a realistic driving study. *IET Intell. Transp. Syst.* **14**(10), 1265–1277 (2020)
94. Raja, M., Exler, A., Hemminki, S., Konomi, S., Sigg, S., Inoue, S.: Towards pervasive geospatial affect perception. *GeoInformatica* **22**(1), 143–169 (2018)
95. Boril, H., Sadjadi, S.O., Hansen, J.H.: Utdrive: Emotion and cognitive load classification for in-vehicle scenarios. In: *Biennial Workshop Digit. Signal Process. In-Vehicle Syst.* (2011)
96. Kamaruddin, N., Wahab, A.: Driver behavior analysis through speech emotion understanding. In: *2010 IEEE Intelligent Vehicles Symposium*, pp. 238–243 (2010). IEEE
97. Nor, N.M., Wahab, A.: Driver identification and driver's emotion verification using KDE and MLP neural networks. In: *Proceeding of the 3rd international conference on information and communication technology for the Moslem World (ICT4M) 2010*, pp. 96–101 (2010). IEEE
98. Lee, S., Lee, T., Yang, T., Yoon, C., Kim, S.-P.: Detection of drivers' anxiety invoked by driving situations using multimodal biosignals. *Processes* **8**(2), 155 (2020)
99. Ooi, J.S.K., Ahmad, S.A., Chong, Y.Z., Ali, S.H.M., Ai, G., Wagatsuma, H.: Driver emotion recognition framework based on electrodermal activity measurements during simulated driving conditions. In: *2016 IEEE EMBS conference on biomedical engineering and sciences (IECBES)*, pp. 365–369 (2016). IEEE
100. Rebollo-Mendez, G., Reyes, A., Paszkowicz, S., Domingo, M.C., Skrypchuk, L.: Developing a body sensor network to detect emotions during driving. *IEEE Trans. Intell. Transp. Syst.* **15**(4), 1850–1854 (2014)
101. Nasoz, F., Lisetti, C.L., Vasilakos, A.V.: Affectively intelligent and adaptive car interfaces. *Inf. Sci.* **180**(20), 3817–3836 (2010)
102. Wan, P., Wu, C., Lin, Y., Ma, X.: Driving Anger States Detection Based on Incremental Association Markov Blanket and Least Square Support Vector Machine. *Discrete Dyn. Nat. Soc* (2019)
103. Wang, X., Liu, Y., Wang, F., Wang, J., Liu, L., Wang, J.: Feature extraction and dynamic identification of drivers' emotions. *Transp. Res. Pt. F-Traffic Psychol. Behav.* **62**, 175–191 (2019)
104. Cevher, D., Zepf, S., Klinger, R.: Towards multimodal emotion recognition in german speech events in cars using transfer learning. *arXiv preprint arXiv:1909.02764* (2019)
105. Lee, B.G., Chong, T.W., Lee, B.L., Park, H.J., Kim, Y.N., Kim, B.: Wearable mobile-based emotional response-monitoring system for drivers. *IEEE T. Hum. Mach. Syst.* **47**(5), 636–649 (2017)
106. Wan, P., Wu, C., Lin, Y., Ma, X.: On-road experimental study on driving anger identification model based on physiological features by ROC curve analysis. *IET Intell. Transp. Syst.* **11**(5), 290–298 (2017)
107. Malta, L., Miyajima, C., Kitaoka, N., Takeda, K.: Analysis of real-world driver's frustration. *IEEE Trans. Intell. Transp. Syst.* **12**(1), 109–118 (2011)
108. Siedlecka, E., Denson, T.F.: Experimental methods for inducing basic emotions: a qualitative review. *Emot. Rev.* **11**(1), 87–97 (2019)
109. Thompson, E.R.: Development and validation of an internationally reliable short-form of the positive and negative affect schedule (panas). *J. Cross-Cult. Psychol.* **38**(2), 227–242 (2007)
110. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* **25**(1), 49–59 (1994)
111. Gross, J.J., Levenson, R.W.: Emotion elicitation using films. *Cogn. Emot.* **9**(1), 87–108 (1995)
112. Healey, J.A., Picard, R.W.: Detecting stress during real-world driving tasks using physiological sensors. *IEEE Trans. Intell. Transport. Syst.* **6**(2), 156–166 (2005)
113. Saurav, S., Gidde, P., Saini, R., Singh, S.: Dual integrated convolutional neural network for real-time facial expression recognition in the wild. *Visual Comput.* **38**(3), 1083–1096 (2022)
114. Hans, A.S.A., Rao, S.: A cnn-lstm based deep neural networks for facial emotion detection in videos. *Int. J. Adv. Signal Image Sci.* **7**(1), 11–20 (2021)
115. Jain, D.K., Shamsolmoali, P., Sehdev, P.: Extended deep neural network for facial emotion recognition. *Pattern Recognit. Lett.* **120**, 69–74 (2019)
116. Liu, C., Tang, T., Lv, K., Wang, M.: Multi-feature based emotion recognition for video clips. In: *Proc. ACM Int. Conf. Multimodal Interact.*, pp. 630–634 (2018)

117. Jain, N., Kumar, S., Kumar, A., Shamsolmoali, P., Zareapoor, M.: Hybrid deep neural networks for face emotion recognition. *Pattern Recognit. Lett.* **115**, 101–106 (2018)
118. Yang, D., Alsadoon, A., Prasad, P.C., Singh, A.K., Elchouemi, A.: An emotion recognition model based on facial recognition in virtual learning environment. *Procedia Comput. Sci.* **125**, 2–10 (2018)
119. Senthilkumar, N., Karpakam, S., Devi, M.G., Balakumaresan, R., Dhilipkumar, P.: Speech emotion recognition based on bi-directional lstm architecture and deep belief networks. *Materials Today: Proc.* **57**, 2180–2184 (2022)
120. Kumaran, U., Radha Rammohan, S., Nagarajan, S.M., Prathik, A.: Fusion of mel and gammatone frequency cepstral coefficients for speech emotion recognition using deep c-rnn. *Int. J. Speech Technol.* **24**(2), 303–314 (2021)
121. Li, Y., Zhao, T., Kawahara, T.: Improved end-to-end speech emotion recognition using self attention mechanism and multitask learning. In: *Interspeech*, pp. 2803–2807 (2019)
122. Albanie, S., Nagrani, A., Vedaldi, A., Zisserman, A.: Emotion recognition in speech using cross-modal transfer in the wild. In: *Proc. ACM Int. Conf. Multimedia*, pp. 292–301 (2018)
123. Tzirakis, P., Trigeorgis, G., Nicolaou, M.A., Schuller, B.W., Zafeiriou, S.: End-to-end multimodal emotion recognition using deep neural networks. *IEEE J. Sel. Topics Signal Process.* **11**(8), 1301–1309 (2017)
124. Shen, F., Peng, Y., Dai, G., Lu, B., Kong, W.: Coupled projection transfer metric learning for cross-session emotion recognition from eeg. *Syst.* **10**(2), 47 (2022)
125. Liu, W., Zheng, W.-L., Li, Z., Wu, S.-Y., Gan, L., Lu, B.-L.: Identifying similarities and differences in emotion recognition with eeg and eye movements among chinese, german, and french people. *J. Neural Eng.* **19**(2), 026012 (2022)
126. Wang, Y., Qiu, S., Li, D., Du, C., Lu, B.-L., He, H.: Multi-modal domain adaptation variational autoencoder for eeg-based emotion recognition. *IEEE/CAA J. Autom. Sinica* (2022)
127. Peng, Y., Zhang, Y., Kong, W., Nie, F., Lu, B.-L., Cichocki, A.: S3lrr: A unified model for joint discriminative subspace identification and semi-supervised eeg emotion recognition. *IEEE Trans. Instrum. Meas.* (2022)
128. Li, X., Shen, F., Peng, Y., Kong, W., Lu, B.-L.: Efficient sample and feature importance mining in semi-supervised eeg emotion recognition. *Express Briefs, IEEE Trans. Circuits Syst. II* (2022)
129. Zhang, G., Yu, M., Liu, Y.-J., Zhao, G., Zhang, D., Zheng, W.: Sparsedgcnn: recognizing emotion from multichannel eeg signals. *IEEE Trans. Affect. Comput.* **14**(1), 537–548 (2021)
130. Lian, Z., Liu, B., Tao, J.: Smin: Semi-supervised multi-modal interaction network for conversational emotion recognition. *IEEE Trans. Affect. Comput.* (2022)
131. Zhang, K., Li, Y., Wang, J., Cambria, E., Li, X.: Real-time video emotion recognition based on reinforcement learning and domain knowledge. *IEEE Trans. Circuits Syst. Video Technol.* (2021)
132. Nie, W., Yan, Y., Song, D., Wang, K.: Multi-modal feature fusion based on multi-layers lstm for video emotion recognition. *Multi-media Tools Appl.* **80**(11), 16205–16214 (2021)
133. Tripathi, S., Tripathi, S., Beigi, H.: Multi-modal emotion recognition on iemocap dataset using deep learning. *arXiv:1804.05788* (2018)
134. Bustos, C., Elhaoui, N., Sole-Ribalta, A., Borge-Holthoer, J., Lapedriza, A., Picard, R.: Predicting driver self-reported stress by analyzing the road scene. In: *2021 9th international conference on affective computing and intelligent interaction (ACII)*, pp. 1–8 (2021). IEEE
135. Liu, S., Koch, K., Zhou, Z., Föll, S., He, X., Menke, T., Fleisch, E., Wortmann, F.: The empathetic car: Exploring emotion inference via driver behaviour and traffic context. *Proc. ACM Interact. Mobile Wearable and Ubiquitous Technol.* **5**(3), 1–34 (2021)
136. Bethge, D., Kosch, T., Grosse-Puppenthal, T., Chuang, L.L., Kari, M., Jagaciak, A., Schmidt, A.: Vemotion: Using driving context for indirect emotion prediction in real-time. In: *Annu. ACM Symp. User Interface Softw. Technol.*, pp. 638–651 (2021)
137. Dobbins, C., Fairclough, S.: Detecting and visualizing context and stress via a fuzzy rule-based system during commuter driving. In: *IEEE Int. Conf. Pervasive Comput. Commun. Workshops*, pp. 499–504 (2019). IEEE
138. Bitkina, O.V., Kim, J., Park, J., Kim, H.K.: Identifying traffic context using driving stress: a longitudinal preliminary case study. *Sensors* **19**(9), 2152 (2019)
139. Ding, N., Ma, H., Zhao, C., Ma, Y., Ge, H.: Driver's emotional state-based data anomaly detection for vehicular ad hoc networks. In: *IEEE Int. Conf. Smart Internet of Things*, pp. 121–126 (2019)
140. Zhang, Y., Chen, M., Guizani, N., Wu, D., Leung, V.C.: Sovcan: safety-oriented vehicular controller area network. *IEEE Commun. Mag.* **55**(8), 94–99 (2017)
141. Vögel, H.-J., Stieß, C., Hubregtsen, T., André, E., Schuller, B., Härrri, J., Conrad, J., Adi, A., Zadorojnyi, A., Terken, J., et al.: Emotion-awareness for intelligent vehicle assistants: A research agenda. In: *IEEE/ACM Int. Workshop Softw. Eng. AI Auton. Syst.*, pp. 11–15 (2018)
142. Raja, M., Sigg, S.: Rfexpress!-exploiting the wireless network edge for rf-based emotion sensing. In: *IEEE Int. Conf. Emerg. Technol. Factory Automat.*, pp. 1–8 (2017)
143. Li, W., Zhang, B., Wang, P., Sun, C., Zeng, G., Tang, Q., Guo, G., Cao, D.: Visual-attribute-based emotion regulation of angry driving behaviours. *IEEE Intell. Transp. Syst. Mag.* **14**(3), 10–28 (2022)
144. Völkel, S.T., Graefe, J., Schödel, R., Häuslschmid, R., Stachel, C., Au, Q., Hussmann, H.: I drive my car and my states drive me: Visualizing driver's emotional and physical states. In: *Proc. Int. Conf. Automot. User Interfaces Interact. Veh. Appl.*, pp. 198–203 (2018)
145. Spiridon, E., Fairclough, S.: The effects of ambient blue light on anger levels: applications in the design of unmanned aircraft gcs. *Int. J. Unmanned Syst. Eng.* **5**(3), 53–69 (2017)
146. Johnson, M.B., McKnight, S.: Warning drivers about potential congestion as a means to reduce frustration-driven aggressive driving. *Traffic Inj. Prev.* **10**(4), 354–360 (2009)
147. Braun, M., Chadowitz, R., Alt, F.: User experience of driver state visualizations: A look at demographics and personalities. In: *Human-Computer Interaction-INTERACT 2019: 17th IFIP TC 13 International Conference, Paphos, Cyprus, September 2–6, 2019, Proceedings, Part IV 17*, pp. 158–176 (2019). Springer
148. Paredes, P.E., Balters, S., Qian, K., Murnane, E.L., Ordóñez, F., Ju, W., Landay, J.A.: Driving with the fishes: Towards calming and mindful virtual reality experiences for the car. *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.* **2**(4), 1–21 (2018)
149. Hassib, M., Braun, M., Pflöging, B., Alt, F.: Detecting and influencing driver emotions using psycho-physiological sensors and ambient light. In: *IFIP Conference on Human-Computer Interaction*, pp. 721–742 (2019). Springer
150. Balters, S., Landay, J.A., Paredes, P.E.: On-road guided slow breathing interventions for car commuters. In: *CHI Conf. Hum. Factors Comput. Syst.*, pp. 1–5 (2019)
151. Balters, S., Mauriello, M.L., Park, S.Y., Landay, J.A., Paredes, P.E.: Calm commute: guided slow breathing for daily stress management in drivers. *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.* **4**(1), 1–19 (2020)
152. Paredes, P.E., Zhou, Y., Hamdan, N.A.-H., Balters, S., Murnane, E., Ju, W., Landay, J.A.: Just breathe: In-car interventions for guided slow breathing. *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.* **2**(1), 1–23 (2018)

153. Balters, S., Murnane, E.L., Landay, J.A., Paredes, P.E.: Breath booster! exploring in-car, fast-paced breathing interventions to enhance driver arousal state. In: Proc. EAI Int. Conf. Pervasive Comput. Technol. Healthcare, pp. 128–137 (2018)
154. Schmidt, E., Decke, R., Rasshofer, R., Bullinger, A.C.: Psychophysiological responses to short-term cooling during a simulated monotonous driving task. *Appl. Ergon.* **62**, 9–18 (2017)
155. Dass Jr, D.E., Uyttendaele, A., Terken, J.: Haptic in-seat feedback for lane departure warning. In: Proc. Int. Conf. Automot. User Interfaces Interact. Veh. Appl., pp. 258–261 (2013)
156. Dmitrenko, D., Maggioni, E., Brianza, G., Holthausen, B.E., Walker, B.N., Obrist, M.: Caroma therapy: pleasant scents promote safer driving, better mood, and improved well-being in angry drivers. In: Proc. CHI Conf. Hum. Factors Comput. Syst., pp. 1–13 (2020)
157. Raudenbush, B., Grayhem, R., Sears, T., Wilson, I.: Effects of peppermint and cinnamon odor administration on simulated driving alertness, mood and workload. *N. Am. J. Psychol.* **11**(2), 245–256 (2009)
158. Yoshida, M., Kato, C., Kawasumi, M., Yamasaki, H., Yamamoto, S., Nakano, T., Yamada, M.: Study on stimulation effects for driver based on fragrance presentation. In: MVA, pp. 332–335 (2011)
159. Dmitrenko, D., Maggioni, E., Obrist, M.: I smell trouble: using multiple scents to convey driving-relevant information. In: Proc. 20th ACM Int. Conf. Multimodal Interact., pp. 234–238 (2018)
160. Dmitrenko, D., Maggioni, E., Vi, C.T., Obrist, M.: What did i sniff? mapping scents onto driving-related messages. In: Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, pp. 154–163 (2017)
161. Dmitrenko, D., Maggioni, E., Obrist, M.: Towards a framework for validating the matching between notifications and scents in olfactory in-car interaction. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–6 (2019)
162. Funato, H., Yoshikawa, M., Kawasumi, M., Yamamoto, S., Yamada, M., Yanagida, Y.: Stimulation effects provided to drivers by fragrance presentation considering olfactory adaptation. In: 2009 IEEE Intelligent Vehicles Symposium, pp. 881–886 (2009). IEEE
163. Mustafa, M., Rustam, N., Siran, R.: The impact of vehicle fragrance on driving performance: What do we know. *Procd. Soc. Behv.* **222**, 807–815 (2016)
164. Jeon, M., Walker, B.N., Gable, T.M.: The effects of social interactions with in-vehicle agents on a driver's anger level, driving performance, situation awareness, and perceived workload. *Applied ergonom.* **50**, 185–199 (2015)
165. Li, S., Zhang, T., Zhang, W., Liu, N., Lyu, G.: Effects of speech-based intervention with positive comments on reduction of driver's anger state and perceived workload, and improvement of driving performance. *Applied Ergonom.* **86**, 103098 (2020)
166. FakhrHosseini, S.M., Jeon, M.: How do angry drivers respond to emotional music? a comprehensive perspective on assessing emotion. *J. Multimodal User Interfaces* **13**(2), 137–150 (2019)
167. Fakhrhosseini, S.M., Landry, S., Tan, Y.Y., Bhattarai, S., Jeon, M.: If you're angry, turn the music on: Music can mitigate anger effects on driving performance. In: Proc. Int. Conf. Automot. User Interfaces Interact. Veh. Appl., pp. 1–7 (2014)
168. FakhrHosseini, M., Jeon, M.: The effects of various music on angry drivers' subjective, behavioral, and physiological states. In: Proc. Int. Conf. Automot. User Interfaces Interact. Veh. Appl., pp. 191–196 (2016)
169. Zhu, Y., Wang, Y., Li, G., Guo, X.: Recognizing and releasing drivers' negative emotions by using music: Evidence from driver anger. In: Proc. Int. Conf. Automot. User Interfaces Interact. Veh. Appl., pp. 173–178 (2016)
170. Braun, M., Schubert, J., Pflöging, B., Alt, F.: Improving driver emotions with affective strategies. *Multimodal Technol. Interact.* **3**(1), 21 (2019)
171. Jonsson, I.-M.: Social and emotional characteristics of speech-based in-vehicle information systems: impact on attitude and driving behaviour. PhD thesis, Linköping University Electronic Press (2009)
172. Harris, H., Nass, C.: Emotion regulation for frustrating driving contexts. In: Proc. SIGCHI Conf. Hum. Factors in Comput. Syst., pp. 749–752 (2011)
173. Oehl, M., Lienhop, M., Ihme, K.: Mitigating frustration in the car: Which emotion regulation strategies might work for different age groups? In: HCI International 2021-Posters: 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part III 23, pp. 273–280 (2021). Springer
174. Pêcher, C., Lemerrier, C., Cellier, J.-M.: Emotions drive attention: effects on driver's behaviour. *Saf. Sci.* **47**(9), 1254–1259 (2009)
175. Nass, C., Jonsson, I.-M., Harris, H., Reaves, B., Endo, J., Brave, S., Takayama, L.: Improving automotive safety by pairing driver emotion and car voice emotion. In: CHI Conf. Hum. Factors Comput. Syst., pp. 1973–1976 (2005)
176. Jonsson, I.-M., Nass, C., Endo, J., Reaves, B., Harris, H., Ta, J.L., Chan, N., Knapp, S.: Don't blame me i am only the driver: Impact of blame attribution on attitudes and attention to driving task. In: CHI Extended Abstr. on Human Factors Comput. Syst., pp. 1219–1222 (2004)
177. Jonsson, I.-M., Zajicek, M., Harris, H., Nass, C.: Thank you, i did not see that: in-car speech based information systems for older adults. In: CHI Extended Abstr. on Human Factors Comput. Syst., pp. 1953–1956 (2005)
178. Brodsky, W., Kizner, M.: Exploring an alternative in-car music background designed for driver safety. *Transp. Res. Pt. F-Traffic Psychol. Behav.* **15**(2), 162–173 (2012)
179. Hsieh, L., Seaman, S., Young, R.: Effect of emotional speech tone on driving from lab to road: fmri and erp studies. In: Proc. Int. Conf. Automot. User Interfaces Interact. Veh. Appl., pp. 22–28 (2010)
180. Van Der Zwaag, M.D., Dijksterhuis, C., De Waard, D., Mulder, B.L., Westerink, J.H., Brookhuis, K.A.: The influence of music on mood and performance while driving. *Ergonomics* **55**(1), 12–22 (2012)
181. Jeon, M., Walker, B.N., Yim, J.-B.: Effects of specific emotions on subjective judgment, driving performance, and perceived workload. *Transp. Res. Pt. F-Traffic Psychol. Behav.* **24**, 197–209 (2014)
182. Oehl, M., Ihme, K., Drewitz, U., Pape, A.-A., Cornelsen, S., Schramm, M.: Towards a frustration-aware assistant for increased in-vehicle ux: F-relacs. In: Proc. Int. Conf. Automot. User Interfaces Interact. Veh. Appl., pp. 260–264 (2019)
183. Hu, X., Deng, J., Zhao, J., et al.: Safedj: A crowd-cloud codesign approach to situation-aware music delivery for drivers. *ACM Trans. Multimed. Comput. Commun.* **12**(1s), 1–24 (2015)
184. Krishnan, A.S., Hu, X., Deng, J.-q., Wang, R., Liang, M., Zhu, C., Leung, V.C., Kwok, Y.-K.: A novel cloud-based crowd sensing approach to context-aware music mood-mapping for drivers. In: IEEE Int. Conf. Cloud Comput. Technol. Sci., pp. 475–478 (2015)
185. Wang, C., Terken, J., Hu, J., Rauterberg, M.: "likes" and "dislikes" on the road: A social feedback system for improving driving behavior. In: Proc. Int. Conf. Automot. User Interfaces Interact. Veh. Appl., pp. 43–50 (2016)
186. Li, W., Xue, J., Tan, R., Wang, C., Deng, Z., Li, S., Guo, G., Cao, D.: Global-local-feature-fused driver speech emotion detection for intelligent cockpit in automated driving. *IEEE Trans. Intell. Vehicles* **8**(4), 2684–2697 (2023)



**Wenbo Li** received the B.S., M.Sc., and Ph.D. degree in Automotive Engineering from Chongqing University, Chongqing, China, in 2014, 2017 and 2021, respectively. From 2018 to 2020, he was a visiting Ph.D. student at University of Waterloo, Waterloo, Canada. He is currently a Postdoctoral Research Fellow at the School of Vehicle and Mobility, Tsinghua University, Beijing, China. His research interests include intelligent vehicle, intelligent cockpit, human emotion and cognition, human-machine

interaction, and brain-computer interface.

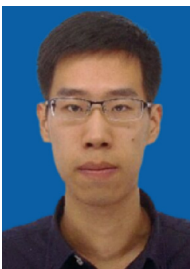


**Guofa Li** received the Ph.D. degree in Mechanical Engineering from Tsinghua University, China, in 2016. He is currently a Professor with Chongqing University, China. His research interests include environment perception, driver behavior analysis, and human-like decision-making based on artificial intelligence technologies in autonomous vehicles and intelligent transportation systems. He has published more than 70 papers in his research areas. He is the recipient of the Young Elite Scientists Sponsorship Program in China, and he receives

the best paper awards from the China Association for Science and Technology (CAST) and the Automotive Innovation journal. In addition, he serves as the Associate Editor for IEEE Transactions on Affective Computing, IEEE Sensors Journal, as well as the lead guest editor for IEEE Intelligent Transportation Systems Magazine and Automotive Innovation.



**Ruichen Tan** received a Bachelor's degree in Mechanical Engineering from Huazhong University of Science and Technology, Wuhan, China, in 2017, and a Master's degree in Electrical and Computer Engineering from Cornell University, Ithaca, NY, USA, in 2019. He is currently pursuing a PhD degree in Civil Engineering with a focus on intelligent transportation systems at Purdue University, West Lafayette, IN, USA. His research interests include Human-related AI, Big Data analysis.

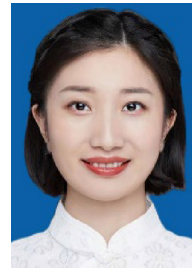


**Cong Wang** received the B.S. and Ph.D. degree in automotive engineering from Beijing Institute of Technology, China, in 2016 and 2022, respectively. He is currently a Postdoctoral Research Fellow with Tsinghua University, Beijing, China. His research interests mainly include vehicle dynamics and safety control for electric vehicles.



**Zemin Sun** received a B.S. degree in Software Engineering, an M.S. degree and a Ph.D. degree in Computer Science and Technology from Jilin University, Changchun, China, in 2015, 2018, and 2022, respectively. She is currently an Assistant Researcher in College of Computer Science and Technology at Jilin University. Her research interests

include mobile edge computing, UAV communications, vehicular networks, and game theory



**Ying Li** received the Ph.D. degree from University of Waterloo, Canada, in 2021. She is currently an Assistant Professor with the School of Mechanical Engineering, Beijing Institute of Technology, China. Her current research focuses on autonomous driving, environmental perception, computer vision, mobile laser scanning, geometric and semantic modeling, and high definition mapping



**Gang Guo** received the Ph.D. degrees in mechanical engineering from Chongqing University, Chongqing, China, in 1994. He is currently the professor at the Department of Mechanical and Vehicle Engineering, Chongqing University. He has authored and coauthored over 100 refereed journal and conference publications. His research interests include human-machine interaction, user experience, intelligent cockpit, intelligent vehicle, and brain computer interface.



**Dongpu Cao** received the Ph.D. degree from Concordia University, Canada, in 2008. He is a Professor at Tsinghua University. His current research focuses on driver cognition, automated driving and cognitive autonomous driving. He has contributed more than 200 papers and 3 books. He received the SAE Arch T. Colwell Merit Award in 2012, IEEE VTS 2020 Best Vehicular Electronics Paper Award and 6 Best Paper Awards from international conferences. He has served as Deputy Editor-in-Chief for IET intelligent

transport systems journal, and an associate editor for IEEE transactions on vehicular technology, IEEE transactions on intelligent transportation systems, IEEE/ASME transactions on mechatronics, IEEE transactions on industrial electronics, IEEE/CAA journal of automatica sinica, IEEE transactions on computational social systems, IEEE transactions on intelligent vehicles, and ASME journal of dynamic systems, measurement and control. He is an IEEE VTS Distinguished Lecturer.



**Keqiang Li** received a B.E. degree from Tsinghua University, Beijing, China, in 1985, and M.E. and Ph.D. degrees from Chongqing University, Chongqing, China, in 1988 and 1995, respectively. He is currently a professor at the School of Vehicle and Mobility, Tsinghua University. His current research interests include intelligent connected vehicles, cloud-based control for vehicles, and vehicle dynamics systems. He is the Chief Scientist of the Intelligent and Connected Vehicle Innovation Center of China, and the

Director of the State Key Laboratory of Automotive Safety and Energy of China.